Purification of quantum trajectories

Hans Maassen Burkhard Kümmerer

Abstract: We prove that the quantum trajectory of repeated perfect measurement on a finite quantum system either asymptotically purifies, or hits upon a family of 'dark' subspaces, where the time evolution is unitary.

1. Introduction

A key concept in the modern theory of open quantum systems is the notion of indirect measurement as introduced by Kraus [Kra]. An *indirect measurement* on a quantum system is a (direct) measurement of some quantity in its environment, made after some interaction with the system has taken place.

When we make such a measurement, our description of the quantum system changes in two ways: we account for the flow of time by a unitary transformation (following Schrödinger), and we update our knowledge of the system by conditioning on the measurement outcome (following von Neumann). If we then repeat the indirect measurement indefinitely, we obtain a chain of random outcomes. In the course of time we may keep record of the updated density matrix Θ_t, which at time t reflects our best estimate of all observable quantities of the quantum system, given the observations made up to that time. This information can in its turn be used to predict later measurements outcomes. The stochastic process Θ_t of updated states, is the *quantum trajectory* associated to the repeated measurement process.

By taking the limit of continuous time, we arrive at the modern models of continuous observation: quantum trajectories in continuous time satisfying *stochastic Schrödinger equations* [Dav], [Gis], [Car], [BGM]. These models are employed with great success for calculations and computer simulations of laboratory experiments such as photon counting and homodyne field detection.

In this paper we consider the question, what happens to the quantum trajectory at large times. We do so only for the case of discrete time, not a serious restriction indeed, since asymptotic behaviour remains basically unaltered in the continuous time limit.

We focus on the case of *perfect measurement*, i.e. the situation where no information flows into the system, and all information which leaks out is indeed observed. In classical probability such repeated perfect measurement would lead to a further and further narrowing of the distribution of the system, until it either becomes pure, i.e. an atomic measure, or it remains spread out over some area, thus leaving a certain amount of information 'in the dark' forever. Using a fundamental inequality of Nielsen [Nie] we prove that in quantum mechanics the situation is quite comparable: the density matrix tends to purify, until it hits upon some family of 'dark' subspaces, if such exist, i.e. spaces from which no information can leak out. A crucial difference with the classical case is, however, that even after all available information has been extracted by observation, the state continues to move about in a random fashion between the 'dark' subspaces, thus continuing to produce 'quantum noise'.

The structure of this paper is as follows. In Section 2 we introduce quantum measurement on a finite system, in particular Kraus measurement. In Section 3 repeated

*This paper is dedicated to Mike Keane on his 65-th birthday

Keywords and phrases: quantum trajectory, stochastic Schrödinger equation

measurement and the quantum trajectory are introduced, and in Section 4 we prove our main result. Some typical examples of dark subspaces are given in Section 5.

2. A single measurement

Let $\mathcal{A}$ be the algebra of all complex $d \times d$ matrices. By $\mathcal{S}$ we denote the space of $d \times d$ *density matrices*, i.e. positive matrices of trace 1. We think of $\mathcal{A}$ as the observable algebra of some finite quantum system, and of $\mathcal{S}$ as the associated state space.
A measurement on this quantum system is an operation which results in the extraction of information while possibly changing its state. Before the measurement the system is described by a *prior* state $\theta \in \mathcal{S}$, and afterwards we obtain a piece of information, say an outcome $i \in \{1, 2, \ldots, k\}$, and the system reaches some new (or *posterior*) state θ_i':

$$\theta \longrightarrow (i, \theta_i') \ .$$

Now, a probabilistic theory, rather than predicting the outcome i, gives a probability distribution $(\pi_1, \pi_2, \ldots, \pi_k)$ on the possible outcomes. Let

$$T_i : \theta \mapsto \pi_i \theta_i' \ , \qquad (i = 1, \ldots, k) \ . \tag{1}$$

Then the operations T_i, which must be completely positive, code for the probabilities $\pi_i = \mathrm{tr}(T_i\theta)$ of the possible outcomes, as well as for the posterior states $\theta_i' = T_i\theta/\mathrm{tr}(T_i\theta)$, conditioned on these outcomes. The k-tuple $(T_1, \ldots, T_k)$ describes the quantum measurement completely. Its mean effect on the system, averaged over all possible outcomes, is given by the trace-preserving map

$$T : \theta \mapsto \sum_{i=1}^{k} \pi_i \theta_i' = \sum_{i=1}^{k} T_i \theta \ .$$

Example 1: *von Neumann measurement.*
Let $p_1, p_2, \ldots p_k$ be mutually orthogonal projections in $\mathcal{A}$ adding up to $\mathbf{1}$, and let $a \in \mathcal{A}$ be a self-adjoint matrix whose eigenspaces are the ranges of the p_i. Then according to von Neumann's projection postulate a measurement of a is obtained by choosing for T_i the operation

$$T_i(\theta) = p_i \theta p_i \ .$$

Example 2: *Kraus measurement.*
The following indirect measurement procedure was introduced by Karl Kraus [Kra]. It is contains von Neumann's measurement as an ingredient, but is considerably more flexible and realistic.
Our quantum system $\mathcal{A}$ in the state θ is brought into contact with a second system, called the 'ancilla', which is described by a matrix algebra $\mathcal{B}$ in the state β. The two systems interact for a while under Schrödinger's evolution, which results in a rotation over a unitary $u \in \mathcal{B} \otimes \mathcal{A}$. Then the ancilla is decoupled again, and is subjected to a von Neumann measurement given by the orthogonal projections $p_1, \ldots, p_k \in \mathcal{B}$. The outcome of this measurent contains information about the system, since system and ancilla have become correlated during their interaction. In order to assess this information, let us consider an event in our quantum system, described by a projection $q \in \mathcal{A}$. Since each of the projections $p_i \otimes \mathbf{1}$ commutes with

$\mathbf{1} \otimes q$, the events of seeing outcome i and then the occurrence of q are compatible, so according to von Neumann we may express the probability for both of them to happen as:

$$\mathbb{P}[\text{outcome } i \text{ and then event } q] = \operatorname{tr} \otimes \operatorname{tr}\Big(\big(u(\beta \otimes \theta)u^*\big)(p_i \otimes q)\Big) .$$

Therefore the following conditional probability makes physical sense.

$$\mathbb{P}[\text{event } q|\text{outcome } i] = \frac{\operatorname{tr} \otimes \operatorname{tr}\Big(\big(u(\beta \otimes \theta)u^*\big)(p_i \otimes q)\Big)}{\operatorname{tr} \otimes \operatorname{tr}\Big(\big(u(\beta \otimes \theta)u^*\big)(p_i \otimes \mathbf{1})\Big)} .$$

This expression, which describes the posterior probability of any event $q \in \mathcal{A}$, can be considered as the posterior state of our quantum system, conditioned on the measurement of an outcome i on the ancilla, even when no event q is subsequently measured. As above, let us therefore call this state θ_i'. We then have

$$\operatorname{tr}(\theta_i' q) = \frac{\operatorname{tr}\big((T_i\theta)q\big)}{\operatorname{tr}(T_i\theta)} ,$$

where $T_i\theta$ takes the form

$$T_i\theta = \operatorname{tr} \otimes \operatorname{id}\Big(\big(u(\beta \otimes \theta)u^*\big)(p_i \otimes \mathbf{1})\Big) .$$

Here, id denotes the identity map $\mathcal{S} \to \mathcal{S}$.
The expression for T_i takes a simple form in the case which will interest us here, namely when the following three conditions are satisfied:

(i) $\mathcal{B}$ consists of all $k \times k$-matrices for some k;
(ii) the orthogonal projections $p_i \in \mathcal{B}$ are one-dimensional (say p_i is the matrix with i-th diagonal entry 1, and all other entries 0);
(iii) β is a pure state (say with state vector $(\beta_1, \ldots, \beta_k) \in \mathbb{C}^k$).

These conditions have the following physical interpretations.

(i) The ancilla is purely quantummechanical;
(ii) the measurement discriminates maximally;
(iii) no new information is fed into the system.

If these conditions are satisfied, u can be written as a $k \times k$ matrix (u_{ij}) of $d \times d$ matrices, and T_i may be written

$$T_i\theta = a_i \theta a_i^* , \tag{2}$$

where

$$a_i = \sum_{j=1}^{k} \beta_j u_{ij} .$$

We note that, by construction,

$$\sum_{i=1}^{k} a_i^* a_i = \sum_{i=1}^{k} \sum_{j=1}^{k} \sum_{j'=1}^{k} \overline{\beta_j} u_{ji}^* u_{ij'} \beta_{j'} = \|\beta\|^2 = \mathbf{1} .$$

This basic rule expresses the preservation of the trace by T.

Definition 1 *By a* perfect measurement *on $\mathcal{A}$ we shall mean a k-tuple $(T_1, \dots, T_k)$ of operations on $\mathcal{S}$, where $T_i\theta$ is of the form $a_i\theta a_i^*$ with $\sum_{i=1}^k a_i^* a_i = \mathbf{1}$.*

Mathematically speaking, the measurement $(T_1, \dots, T_k)$ is perfect iff the Stinespring decomposition of each T_i consists of a single term.
We note that every perfect Kraus measurement is a perfect measurement in the above sense, and that every perfect measurement can be obtained as the result of a perfect Kraus measurement.

3. Repeated measurement

By repeating a measurement on the quantum system $\mathcal{A}$ indefinitely, we obtain a Markov chain with values in the state space $\mathcal{S}$. This is the quantum trajectory which we study in this paper.
Let Ω be the space of infinite outcome sequences $\omega = \{\omega_1, \omega_2, \omega_3, \dots\}$, with $\omega_j \in \{1, \dots, k\}$, and let for $m \in \mathbb{N}$ and $i_1, \dots, i_m \in \{1, \dots, k\}$ the *cylinder set* $\Lambda_{i_1,\dots,i_m} \subset \Omega$ be given by

$$\Lambda_{i_1,\dots,i_m} := \{\omega \in \Omega \,|\, \omega_1 = i_1, \dots, \omega_m = i_m\} \,.$$

Denote by Σ_m the Boolean algebra generated by these cylinder sets, and by Σ the σ-algebra generated by all these Σ_m. Let $T_1, \dots, T_k$ be as in Section 2.
Then for every initial state θ_0 on $\mathcal{A}$ there exists a unique probability measure $\mathbb{P}_{\theta_0}$ on (Ω, Σ) satisfying

$$\mathbb{P}_{\theta_0}(\Lambda_{i_1,\dots,i_m}) = \mathrm{tr}(T_{i_m} \circ \dots \circ T_{i_1}(\theta_0)) \,.$$

Indeed, according to the Kolmogorov-Daniell reconstruction theorem we only need to check consistency: since $T = \sum_{i=1}^k T_i$ preserves the trace,

$$\begin{aligned} \sum_{i=1}^k \mathbb{P}_{\theta_0}(\Lambda_{i_1,\dots,i_m,i}) &= \sum_{i=1}^k \mathrm{tr}\big(T_i \circ T_{i_m} \circ \dots \circ T_{i_1}(\theta_0)\big) = \mathrm{tr}(T \circ T_{i_m} \circ \dots \circ T_{i_1}(\theta_0)) \\ &= \mathrm{tr}\big(T_{i_m} \circ \dots \circ T_{i_1}(\theta_0)\big) = \mathbb{P}_{\theta_0}\big(\Lambda_{i_1,\dots,i_m}\big) \,. \end{aligned}$$

On the probability space $(\Omega, \Sigma, \mathbb{P}_{\theta_0})$ we now define the *quantum trajectory* $(\Theta_n)_{n\in\mathbb{N}}$ as the sequence of random variables given by

$$\Theta_n : \Omega \to \mathcal{S} : \omega \mapsto \frac{T_{\omega_m} \circ \dots \circ T_{\omega_1}(\theta_0)}{\mathrm{tr}\big(T_{\omega_m} \circ \dots \circ T_{\omega_1}(\theta_0)\big)} \,.$$

We note that Θ_n is Σ_n-measurable. The density matrix $\Theta_n(\omega)$ describes the state of the system at time n under the condition that the outcomes $\omega_1, \dots, \omega_n$ have been seen.
The quantum trajectory $(\Theta_n)_{n\in\mathbb{N}}$ is a Markov chain with transitions

$$\theta \longrightarrow \theta_i' = \frac{T_i\theta}{\mathrm{tr}(T_i\theta)} \qquad \text{with probability} \quad \mathrm{tr}(T_i\theta) \,. \tag{3}$$

4. Purification

In a perfect measurement, when T_i is of the form $\theta \mapsto a_i \theta a_i^*$, a pure prior state $\theta = |\psi\rangle\langle\psi|$ leads to a pure posterior state:

$$\theta_i' = \frac{a_i|\psi\rangle\langle\psi|a_i^*}{\langle\psi, a_i^* a_i \psi\rangle} = |\psi_i\rangle\langle\psi_i| \ , \quad \text{where} \quad \psi_i = \frac{a_i \psi}{\|a_i \psi\|} \ .$$

Hence in the above Markov chain the pure states form a closed set. Experience with quantum trajectories leads one to believe that in many cases even more is true: along a typical trajectory the density matrix tends to purify: its spectrum approaches the set $\{0, 1\}$. In Markov chain jargon: the pure states form an asymptotically stable set.

There is, however, an obvious counterexample to this statement in general. If every a_i is proportional to a unitary, say $a_i = \sqrt{\lambda_i} u_i$ with $u_i^* u_i = \mathbf{1}$, then

$$\theta_i' = \frac{a_i \theta a_i^*}{\mathrm{tr}(a_i \theta a_i^*)} = u_i \theta u_i^* \sim \theta \ ,$$

where $\sim$ denotes unitary equivalence. So in this case the eigenvalues of the density matrix remain unchanged along the trajectory: pure states remain pure and mixed states remain mixed with unchanging weights. In this section we shall show that in dimension 2 this is actually the only exception. (Cf. Corollary 2.) In higher dimensions the situation is more complicated: if the state does not purify, the a_i must be proportional to unitaries on a certain collection of 'dark' subspaces, which they must map into each other. (Cf. Corollary 8.)

In order to study purification we shall consider the *moments* of Θ_n. By the m-th *moment* of a density matrix $\theta \in \mathcal{S}$ we mean $\mathrm{tr}\,(\theta^m)$. We note that two states θ and ρ are unitarily equivalent iff all their moments are equal. In dimension d equality of the moments $m = 1, \ldots, d$ suffices.

Definition 2 *We say that the quantum trajectory* $\big(\Theta_n(\omega)\big)_{n\in\mathbb{N}}$ purifies *when*

$$\forall_{m\in\mathbb{N}} : \quad \lim_{n\to\infty} \mathrm{tr}\big(\Theta_n(\omega)^m\big) = 1 \ .$$

Note that the only density matrices ρ satisfying $\mathrm{tr}(\rho^m) = 1$ are one-dimensional projections, the density matrices of pure states. In fact, it suffices that the second moment be equal to 1.

We now state our main result concerning repeated perfect measurement.

Theorem 1 *Let* $\big(\Theta_n\big)_{n\in\mathbb{N}}$ *be the Markov chain with initial state* θ_0 *and transition probabilities (3). Then one of the following alternatives holds.*

(i) The paths of $(\Theta_n)_{n\in\mathbb{N}}$ *(the quantum trajectories) purify with probability 1, or:*
(ii) there exists a projection $p \in \mathcal{A}$ *of dimension at least two such that*

$$\forall_{i\in\{1,\ldots,k\}} \exists_{\lambda_i \geq 0} : \quad p a_i^* a_i p = \lambda_i p \ .$$

Condition (ii) says that a_i is proportional to an isometry in restriction to the range of p. Note that this condition trivially holds if p is one-dimensional.

Corollary 2 *In dimension* $d = 2$ *the quantum trajectory of a repeated perfect measurement either purifies with probability 1, or all the* a_i*'s are proportional to unitaries.*

If the a_i are all proportional to unitaries, the coupling to the environment is *essentially commutative* in the sense of [KüM].

Our proof starts from an inequality of Michael Nielsen [Nie] to the effect that for all $m \in \mathbb{N}$ and all states θ:

$$\sum_{i=1}^{k} \pi_i \,\mathrm{tr}\big((\theta_i')^m\big) \geq \mathrm{tr}(\theta^m) ,$$

where

$$\pi_i := \mathrm{tr}(a_i \theta a_i^*) \quad \text{and} \quad \theta_i' := \frac{a_i \theta a_i^*}{\mathrm{tr}(a_i \theta a_i^*)} .$$

Nielsen's inequality says that the expected m-th moment of the posterior state is as least as large as the m-th moment of the prior state. In terms of the associated Markov chain we may express this inequality as

$$\forall_{m,n\in\mathbb{N}} : \qquad \mathbb{E}\left(\mathrm{tr}\left(\Theta_{n+1}^m\right) \middle| \Sigma_n\right) \geq \mathrm{tr}\left(\Theta_n^m\right) ,$$

i.e. the moments $M_n^{(m)} := \mathrm{tr}\left(\Theta_n^m\right)_{n\in\mathbb{N}}$ are submartingales. Clearly all moments take values in $[0,1]$. Therefore, by the martingale convergence theorem they must converge almost surely to some random variables $M^{(m)}$.

This suggests the following line of proof for our theorem: Since the moments converge, the eigenvalues of $(\Theta_n)_{n\in\mathbb{N}}$ must converge. Hence along a single trajectory the states eventually become unitarily equivalent, i.e. eventually

$$\forall_i : \quad \Theta_n(\omega) \sim \frac{a_i \Theta_n(\omega) a_i^*}{\mathrm{tr}(a_i \Theta_n(\omega) a_i^*)} .$$

But this seems to imply that either Θ_n purifies almost surely, or the a_i's are unitary on the support of Θ_n.

In the following proof of Theorem 1 we shall make this suggestion mathematically precise.

Lemma 3 *In the situation of Theorem 1 one of the following alternatives holds.*

(i) For all $m \in \mathbb{N}$: $\lim_{n\to\infty} \mathrm{tr}\left(\Theta_n^m\right) = 1$ almost surely;

(ii) there exists a mixed state $\rho \in \mathcal{S}$ such that

$$\forall_{i=1,\ldots,k} \exists_{\lambda_i \geq 0} : \quad a_i \rho a_i^* \sim \lambda_i \rho .$$

Proof. For each $m \in \mathbb{N}$ we consider the continuous function

$$\delta_m : \mathcal{S} \to [0,\infty) : \theta \mapsto \sum_{i=1}^{k} \mathrm{tr}(a_i \theta a_i^*) \left(\mathrm{tr}\left(\left(\frac{a_i \theta a_i^*}{\mathrm{tr}(a_i \theta a_i^*)}\right)^m\right) - \mathrm{tr}(\theta^m)\right)^2 .$$

Then, using (2) and (3),

$$\delta_m(\Theta_n) = \mathbb{E}\left(\left(M_{n+1}^{(m)} - M_n^{(m)}\right)^2 \middle| \Sigma_n\right) .$$

Since $\left(M_n^{(m)}\right)_{n\in\mathbb{N}}$ is a positive submartingale bounded by 1, its increments must be square summable:

$$\forall_{m\in\mathbb{N}} : \qquad \sum_{n=0}^{\infty} \mathbb{E}\big(\delta_m(\Theta_n)\big) \leq 1 .$$

In particular

$$\lim_{n\to\infty} \sum_{m=1}^{d} \mathbb{E}\big(\delta_m(\Theta_n)\big) = 0 \,. \tag{4}$$

Now let us assume that (i) is not the case, i.e. for some (and hence for all) $m \geq 2$ the expectation $\mathbb{E}(M^{(m)}) =: \mu_m$ is strictly less than 1. For any $n \in \mathbb{N}$ consider the event

$$A_n := \left\{ \omega \in \Omega \;\middle|\; M_n^{(2)} \leq \frac{\mu_2 + 1}{2} \right\} .$$

Then, since $\mathbb{E}\left(M_n^{(2)}\right)$ is increasing in n, we have for all $n \in \mathbb{N}$:

$$\mathbb{E}(M_n^{(2)}) \leq \mathbb{E}(M^{(2)}) = \mu_2 < 1 \,.$$

Therefore for all $n \in \mathbb{N}$,

$$\begin{aligned} \mu_2 &\geq \mathbb{E}\left(M_n^{(2)} \cdot 1_{[M_n^{(2)} > \frac{\mu_2+1}{2}]}\right) \\ &\geq \frac{\mu_2+1}{2}\, \mathbb{P}\left[M_n^{(2)} > \frac{\mu_2+1}{2}\right] \\ &= \frac{\mu_2+1}{2}\left(1 - \mathbb{P}(A_n)\right) \,, \end{aligned}$$

so that

$$\mathbb{P}(A_n) \geq \frac{1-\mu_2}{1+\mu_2} \,. \tag{5}$$

On the other hand, A_n is Σ_n-measureable and therefore it is a union of sets of the from $\Lambda_{i_1,\ldots,i_n}$. Since Θ_n is Σ_n-measureable, Θ_n is constant on such sets; let us call the constant $\Theta_n(i_1,\ldots,i_n)$. We have the following inequality:

$$\frac{1}{\mathbb{P}(A_n)} \sum_{\Lambda_{i_1,\ldots,i_n} \subset A_n} \mathbb{P}\left(\Lambda_{i_1,\ldots,i_n}\right) \left(\sum_{m=1}^{d} \delta_m\big(\Theta_n(i_1,\ldots,i_n)\big) \right) \leq \frac{1}{\mathbb{P}(A_n)} \sum_{m=1}^{d} \mathbb{E}\big(\delta_m(\Theta_n)\big) \,.$$

On the left hand side we have an average of numbers which are each of the form $\sum_{m=1}^{d} \delta_m(\Theta_n(i_1,\ldots,i_n))$, hence we can choose $(i_1,\ldots,i_n)$ such that $\rho_n := \Theta_n(i_1,\ldots,i_n)$ satisfies, by (5),

$$\sum_{m=1}^{d} \delta_m(\rho_n) \leq \frac{\mu_2+1}{\mu_2-1} \sum_{m=1}^{d} \mathbb{E}\big(\delta_m(\Theta_n)\big) \,.$$

Since $\Lambda_{i_1,\ldots,i_n} \subset A_n$, the sequence $(\rho_n)_{n\in\mathbb{N}}$ lies entirely in the compact set

$$\left\{ \theta \in \mathcal{S} \;\middle|\; \mathrm{tr}(\theta^2) \leq \frac{\mu_2+1}{2} \right\} .$$

Let ρ be a cluster point of this sequence. Then, since $\mathbb{E}(\delta_m(\Theta_n))$ tends to 0 as $n \to \infty$, and δ_m is continuous, we may conclude that for $m = 1,\ldots,d$:

$$\delta_m(\rho) = 0 \,, \quad \text{and} \quad \mathrm{tr}(\rho^2) \leq \frac{\mu_2+1}{2} < 1 \,.$$

So ρ is a mixed state, and, by the definition of δ_m,

$$\operatorname{tr}(a_i\rho a_i^*)\left(\operatorname{tr}\left(\left(\frac{a_i\rho a_i^*}{\operatorname{tr}(a_i\rho a_i^*)}\right)^m\right)-\operatorname{tr}(\rho^m)\right)^2=0$$

for all $m=1,2,3,\ldots,d$ and all $i=1,\ldots,k$. Therefore either $\operatorname{tr}(a_i\rho a_i^*)=0$, i.e. $a_i\rho a_i^*=0$, proving our statement (ii) with $\lambda_i=0$; or $\operatorname{tr}(a_i\rho a_i^*)>0$, in which case $\rho_i':=a_i\rho a_i^*/\operatorname{tr}(a_i\rho a_i^*)$ and ρ itself have the same moments of orders $m=1,2,\ldots,d$, so that they are unitarily equivalent. This proves (ii).

□

From Lemma 3 to Theorem 1 is an exercise in linear algebra:

Lemma 4 *Let $a_1,\ldots,a_k\in M_d$ be such that $\sum_{i=1}^k a_i^*a_i=\mathbf{1}$. Suppose that there exists a density matrix $\rho\in M_d$ such that for $i=1,\ldots,k$*

$$a_i\rho a_i^*\sim\lambda_i\rho\ .$$

Let p denote the support of ρ. Then for all $i=1,\ldots,k$:

$$pa_i^*a_ip=\lambda_ip\ .$$

Proof. Let us define, for a nonnegative matrix x, the *positive determinant* $\det_{\mathrm{pos}}(x)$ to be the product of all its strictly positive eigenvalues (counted with their multiplicities). Then, if p denotes the support projection of x, we have the implication

$$\det\nolimits_{\mathrm{pos}}(x)=\det\nolimits_{\mathrm{pos}}(\lambda p)\quad\Longrightarrow\quad\operatorname{tr}(xp)\geq\operatorname{tr}(\lambda p)\tag{6}$$

with equality iff $x=\lambda p$. (This follows from the fact that the sum of a set of positive numbers with given product is minimal iff these numbers are equal.)
Now let p be the support of ρ as in the Lemma. Let $v_i\sqrt{pa_i^*a_ip}$ denote the polar decomposition of a_ip. Then we have by assumption,

$$\begin{aligned}\det\nolimits_{\mathrm{pos}}(\lambda_i\rho)&=\det\nolimits_{\mathrm{pos}}(a_i\rho a_i^*)\\&=\det\nolimits_{\mathrm{pos}}(a_ip\rho pa_i^*)\\&=\det\nolimits_{\mathrm{pos}}(v_i\sqrt{pa_i^*a_ip}\rho\sqrt{pa_i^*a_ip}v_i^*)\\&=\det\nolimits_{\mathrm{pos}}(\sqrt{pa_i^*a_ip}\rho\sqrt{pa_i^*a_ip})\\&=\det\nolimits_{\mathrm{pos}}(pa_i^*a_ip)\det\nolimits_{\mathrm{pos}}(\rho)\ .\end{aligned}$$

Now, since $\det_{\mathrm{pos}}(\lambda_i\rho)=\det_{\mathrm{pos}}(\lambda_ip)\cdot\det_{\mathrm{pos}}(\rho)$ and $\det_{\mathrm{pos}}(\rho)>0$, it follows that

$$\det\nolimits_{\mathrm{pos}}(\lambda_ip)=\det\nolimits_{\mathrm{pos}}(pa_i^*a_ip)\ .$$

By the implication (6) we may conclude that

$$\operatorname{tr}(pa_i^*a_ip)\geq\operatorname{tr}\lambda_ip)\ .\tag{7}$$

On the other hand,

$$\sum_{i=1}^k\lambda_i=\sum_{i=1}^k\operatorname{tr}(\lambda_i\rho)=\sum_{i=1}^k\operatorname{tr}(a_i\rho a_i^*)=\operatorname{tr}\left(\rho\left(\sum_{i=1}^k a_ia_i^*\right)\right)=\operatorname{tr}\rho=1\ ,$$

where in the second equality sign the assumption was used again. Then, by (7),

$$\operatorname{tr} p = \sum_{i=1}^{k} \operatorname{tr}(pa_i^* a_i p) \geq \sum_{i=1}^{k} \operatorname{tr}(\lambda_i p) = \left(\sum_{i=1}^{k} \lambda_i \right) \operatorname{tr} p = \operatorname{tr} p \; .$$

So apparently, in this chain, we have equality. But then, since equality is reached in (6), we find that

$$pa_i^* a_i p = \lambda_i p \; .$$

□

5. Dark subspaces

By considering more than one step at a time the following stronger conclusion can be drawn.

Corollary 5 *In the situation of Theorem 1, either the quantum trajectory purifies with probability 1 or there exists a projection p of dimension at least 2 such that for all $l \in \mathbb{N}$ and all $i_1, \ldots, i_l$ there is $\lambda_{i_1,\ldots,i_l} \geq 0$ with*

$$pa_{i_1}^* \cdots a_{i_l}^* a_{i_l} \cdots a_{i_1} p = \lambda_{i_1,\ldots,i_l} p \; . \tag{8}$$

We shall call a projection p satisfying (8) a *dark* projection, and its range a *dark* subspace.
Let p be a dark projection, and let $v_i \sqrt{pa_i^* a_i} = \sqrt{\lambda_i} v_i p$ be the polar decomposition of $a_i p$. Then the projection $p_i' := v_i p v_i^*$ satisfies:

$$\begin{aligned} \lambda_i p_i' a_{i_1}^* \cdots a_{i_m}^* a_{i_m} \cdots a_{i_1} p_i' &= \lambda_i (v_i p v_i^*) a_{i_1}^* \cdots a_{i_m}^* a_{i_m} \cdots a_{i_1} (v_i p v_i^*) \\ &= v_i p a_i^* a_{i_1}^* \cdots a_{i_m}^* a_{i_m} \cdots a_{i_1} a_i p v_i^* \\ &= \lambda_{i,i_1,\ldots,i_m} \cdot p_i' \; . \end{aligned}$$

Hence if p is dark, and $\lambda_i \neq 0$ then also p_i' is dark with constants

$$\lambda'_{i_1,\ldots,i_m} = \lambda_{i,i_1,\ldots,i_m} / \lambda_i \; .$$

We conclude that asymptotically the quantum trajectory performs a random walk between dark subspaces of the same dimension, with transition probabilities $p \longrightarrow p_i'$ equal to λ_i, the scalar value in $pa_i^* a_i p = \lambda_i p$. In the trivial case that the dimension of p is 1, purification has occurred.
Inspection of the a_i should reveal the existence of nontrivial dark subspaces. If none exist, then purification is certain.
We end this Section with two examples where nontrivial dark subspaces occur.

Example 1. Let $d = l \cdot e$ and let $\mathcal{H}_1, \ldots, \mathcal{H}_l$ be mutually orthogonal e-dimensional subspaces of $\mathcal{H} = \mathbb{C}^d$. Let (π_{ij}) be an $l \times l$ matrix of transition probabilities. Define $a_{ij} \in \mathcal{A}$ by

$$a_{ij} := \sqrt{\pi_{ij}} v_{ij} \; ,$$

where the maps $v_{ij} : \mathcal{H}_i \to \mathcal{H}_j$ are isometric. Then the matrices a_{ij}, $i, j = 1, \ldots, l$ define a perfect measurement whose dark subspaces are $\mathcal{H}_1, \ldots, \mathcal{H}_l$.

Example 2.
The following example makes clear that nontrivial dark subspaces need not be orthogonal.
Let $\mathcal{H} = \mathbb{C}^2 \otimes \mathcal{D}$, where $\mathcal{D}$ is some finite dimensional Hilbert space, and for $i = 1, \ldots, k$ let $a_i := b_i \otimes u_i$, where the 2×2-matrices b_i satisfy the usual equality

$$\sum_{i=1}^{k} b_i^* b_i = \mathbf{1} \ ,$$

and the u_i are unitaries $\mathcal{D} \to \mathcal{D}$. Suppose that the b_i are not all proportional to unitaries. Then the quantum trajectory defined by the a_i has dark subspaces $\psi \otimes \mathcal{D}$, with ψ running through the unit vectors in $\mathbb{C}^2$. Physically this example describes a pair of systems without any interaction between them, one of which is coupled to the environment in an essentially commutative way, whereas the other purifies.

References

[BGM] L. Bouten, M. Guţă, H. Maassen: *Stochastics Schrödinger equations*, Optics and Spectroscopy 94 (2003) 911-919.

[Car] H.J. Carmichael: *An Open systems Approach to Quantum Optics*, Springer-Verlag, Berlin Heidelberg, New York 1993.

[Dav] E.B. Davies: *Quantum Theory of Open Systems*, Academic Press, London, New York, San Francisco 1976.

[Gis] N. Gisin: *Quantum measurements and stochastic processes*, Phys. Rev. Lett. **52** (1984) 1657–1660.

[Kra] Karl Kraus: *States, effects and operations: fundamental notions of quantum theory*, Lecture Notes in Physics, Vol. 190. Springer-Verlag, Berlin 1983.

[KüM] Burkhard Kümmerer, Hans Maassen: *The essentially commutative dilations of dynamical semigroups on M_n*, Commun. Math. Physics **109** (1987) 1–27.

[Nie] Michael Nielsen: *Characterizing mixing and measurement in quantum mechanics*, Physical Review A **63** (2000) art. nr. 022114.

Radboud University Nijmegen
e-mail: `maassen@math.ru.nl`

Technische Universität Darmstadt
e-mail: `kuemmerer@mathematik.tu-darmstadt.de`

Information Transfer Implies State Collapse

B. Janssens and H. Maassen
Mathematisch Instituut, Radboud Universiteit Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands,
maassen@math.kun.nl, basjanss@sci.kun.nl
(February 16, 2006)

We attempt to clarify certain puzzles concerning state collapse and decoherence. In open quantum systems decoherence is shown to be a necessary consequence of the transfer of information to the outside; we prove an upper bound for the amount of coherence which can survive such a transfer. We claim that in large closed systems decoherence has never been observed, but we will show that it is usually harmless to assume its occurrence. An independent postulate of state collapse over and above Schrödinger's equation and the probability interpretation of quantum states, is shown to be redundant.

PACS number: 03.65.Ta

I. INTRODUCTION

In its most basic formulation, quantum theory encodes the preparation of a system in a pure quantum state, a unit vector ψ in a Hilbert space $\mathcal{H}$. Observables are modelled by (say, nondegenerate) self-adjoint operators on $\mathcal{H}$. The expectation value of an observable A in a state ψ is given by $\langle\psi, A\psi\rangle$. If a is an eigenvalue of A and ψ_a a unit eigenvector, and information concerning A is somehow extracted from the system, then the probability for the value a to be observed is $|\langle\psi_a, \psi\rangle|^2$. If this observation is indeed made, then the subsequent behaviour of the system is predicted using the pure state ψ_a. This is called *state collapse*. It follows that, if the information extraction has taken place but the information on the value of A is disregarded, then the subsequent behaviour can be described optimally using a mixture of eigenstates. This is called *decoherence*. In this paper we substantiate the following claim concerning decoherence and state collapse.

Decoherence is only observed in open systems, where it is a necessary consequence of the transfer of information to the outside.

So the observed occurrence of decoherence does not contradict the unitary time evolution postulated by quantum mechanics, since open systems do not evolve unitarily. Decoherence can be explained in quantum theory by embedding the quantum system into a larger, closed whole, which in itself evolves unitarily. This is well-known (see e.g. [Neu]). We add the observation that decoherence is not only a *possibility* for an open system, but a *necessary consequence* of the leakage of information out of the system. We prove an inequality relating the decoherence between two pure states to the degree in which a decision between the two is possible by a measurement outside. This is the content of Theorem 3 in section III.

Also, we have claimed that one *does not* actually observe decoherence in closed macroscopic systems. First of all, most of the systems that are ever observed are actually open, since it is extremely difficult to shield large systems from interaction. But more to the point, the difference between coherence and decoherence can only be seen by measuring some highly exotic 'stray observables' which are almost always forbiddingly hard to observe. And indeed, in those rare cases where experimenters have succeeded in measuring them, ordinary unitary evolution was found, not decoherence. (See [Arn], [Fri], [Wal].)

We illustrate the latter point in section IV, where we show that the measurement of two classes of observables *can not* reveal the difference between coherence and decoherence: a class of microscopic observables and a large class of macroscopic observables. Take as an example a volume of gas. Microscopic observables such as the position of one particular atom in a gas, only relate to a small fraction of the system. Macroscopic observables like the center of mass of the gas, are the average over a large number of microscopic observables. Belonging neither to the macroscopic nor to the microscopic class, the 'stray observables' referred to above describe detailed correlations between large numbers of atoms in the gas. This kind of information is experimentally almost inaccessible.

Driving home our point concerning decoherence in closed systems: coherent superpositions of macroscopically distinguishable states are not the strange monsters produced by a quantum theory applied outside its domain. They are, on the contrary, everyday occurrences which, however, *can not* be distinguished from the more classical incoherent superpositions in practice, and can therefore always be regarded as such.

II. ABSTRACT INFORMATION EXTRACTION

Quantum phenomena are inherently stochastic. This means that, if quantum systems are prepared in identical ways, then nevertheless different events may be observed. A quantum state describes an ensemble of physical systems, e.g. a beam of particles, and is modelled by a normalized trace-class operator ρ on the Hilbert space. The expectation value of an observable A in the state ρ is then $\mathrm{tr}\,(\rho A)$.

An information extraction or measurement on a quantum state is to be considered as the partition of such an ensemble into subensembles, each subensemble corresponding to a measurement outcome. Let us, in the present section, not wonder *how* the splitting of ensembles can be described by quantum theory, but let us see what such an information extraction, if it can be done, will entail for the subsequent behaviour of the subensembles. Note that this process may serve as part of the preparation for further experiments on the system, so that it must again lead to a state.

A. Information Extraction

For simplicity let us assume that only two outcomes can occur, labelled 0 and 1, say with probabilities p_0 and p_1. The ensemble is then split in two parts, described by their respective states ρ_0 and ρ_1. The map

$$M : \rho \mapsto p_0\rho_0 \oplus p_1\rho_1 \tag{2.1}$$

must be normalized, affine and positive. Indeed, normalization is the property that $p_0 + p_1 = 1$, and positivity is the requirement that states must be mapped to states. The affine property entails that for all states ρ and θ on the original system, and for all $\lambda \in [0, 1]$,

$$M(\lambda\rho + (1-\lambda)\vartheta) = \lambda M(\rho) + (1-\lambda)M(\vartheta)\ .$$

This follows from the physical principle that a system which is prepared in the state ρ with probability λ and in the state ϑ with probability $1-\lambda$, say by tossing a coin, can not be distinguished from a physical system in the state $\lambda\rho+(1-\lambda)\vartheta$. We emphasize that indeed this is a *physical* principle, not a matter of definitions. It states, for instance, that a bundle of particles having 50% spin up and 50% spin down can not be distinguished from a bundle having 50% spin left and 50% spin right. This is a falsifyable statement.

B. State Collapse

The above elementary observations are sufficient to prove that information extraction implies state collapse. If M distinguishes perfectly between the pure states ψ_0 and ψ_1, then of course $p_0 = 1$ in case $\rho = |\psi_0\rangle\langle\psi_0|$, and $p_1 = 1$ if $\rho = |\psi_1\rangle\langle\psi_1|$.

Proposition 1 *Let $\mathcal{T}(\mathcal{H})$ denote the space of trace class operators on a Hilbert space $\mathcal{H}$, and let the map $M : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{H}) \oplus \mathcal{T}(\mathcal{H}) : \rho \mapsto M_0(\rho) \oplus M_1(\rho)$ be the linear extension of some normalized, affine and positive map on the states. Suppose that unit vectors ψ_0 and ψ_1 exist such that*

$$M\big(|\psi_0\rangle\langle\psi_0|\big) = M_0\big(|\psi_0\rangle\langle\psi_0|\big) \oplus 0 \qquad \textit{and} \qquad M\big(|\psi_1\rangle\langle\psi_1|\big) = 0 \oplus M_1\big(|\psi_1\rangle\langle\psi_1|\big)\ . \tag{2.2}$$

Then we have $M\big(|\psi_0\rangle\langle\psi_1|\big) = M\big(|\psi_1\rangle\langle\psi_0|\big) = 0$.

Proof. The positivity of M yields $M\big(|\varepsilon e^{i\varphi}\psi_0 + \psi_1\rangle\langle\varepsilon e^{i\varphi}\psi_0 + \psi_1|\big) \geq 0$ as an operator inequality. In particular, the 0-th component must be positive. As $M_0(|\psi_1\rangle\langle\psi_1|) = 0$, it follows that for all $\varepsilon, \phi \in \mathbb{R}$, we have $\varepsilon^2 M_0\big(|\psi_0\rangle\langle\psi_0|\big) + \varepsilon\left(e^{i\varphi}M_0\big(|\psi_0\rangle\langle\psi_1|\big) + e^{-i\varphi}M_0\big(|\psi_1\rangle\langle\psi_0|\big)\right) \geq 0$. Taking the limit $\varepsilon \downarrow 0$ yields $\left(e^{i\varphi}M_0\big(|\psi_0\rangle\langle\psi_1|\big) + e^{-i\varphi}M_0\big(|\psi_1\rangle\langle\psi_0|\big)\right) \geq 0$ for all $\varphi \in \mathbb{R}$. In particular for $\varphi = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$, implying $M_0\big(|\psi_0\rangle\langle\psi_1|\big) = M_0\big(|\psi_1\rangle\langle\psi_0|\big) = 0$.

Exchanging the roles of ψ_0 and ψ_1 in the argument above results in $M_1\big(|\psi_0\rangle\langle\psi_1|\big) = M_1\big(|\psi_1\rangle\langle\psi_0|\big) = 0$, proving the proposition. □

We may draw two conclusions from Proposition 1. The first is that, for all $|\psi\rangle = \alpha_0|\psi_0\rangle + \alpha_1|\psi_1\rangle$, we have

$$(M_0 + M_1)\big(|\psi\rangle\langle\psi|\big) = (M_0 + M_1)\big(|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|\big) . \tag{2.3}$$

In words: for the prediction of events *after* the splitting of the ensemble in two, it no longer matters whether *before* the splitting the system was in the pure state $|\alpha_0\psi_0+\alpha_1\psi_1\rangle\langle\alpha_0\psi_0+\alpha_1\psi_1|$ or in the mixed state $|\alpha_0|^2|\psi_0\rangle\langle\psi_0|+|\alpha_1|^2|\psi_1\rangle\langle\psi_1|$. This phenomenon, which is a direct consequence of the structure (2.1) of the measurement process, we will call *decoherence*.

The second conclusion from Proposition 1 is the following. For all $|\psi\rangle = \alpha_0|\psi_0\rangle + \alpha_1|\psi_1\rangle$, we have

$$M\big(|\psi\rangle\langle\psi|\big) = |\alpha_0|^2 M_0\big(|\psi_0\rangle\langle\psi_0|\big) \oplus |\alpha_1|^2 M_1\big(|\psi_1\rangle\langle\psi_1|\big) . \tag{2.4}$$

In words: if an ensemble is split in two parts, then the '0-ensemble' will further behave as if the system had been in state ψ_0 instead of ψ prior to splitting, and the '1-ensemble' as if it had been in state ψ_1 instead of ψ. This phenomenon will be called *collapse*.

Throughout this article, we will maintain a sharp distinction between the collapse $M : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{H}) \oplus \mathcal{T}(\mathcal{H})$ and the decoherence $(M_0 + M_1) : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{H})$. The former represents the splitting of an ensemble in two parts by means of measurement, whereas the latter represents the splitting and subsequent recombination of this ensemble.

III. OPEN SYSTEMS

A decoherence-mapping $(M_0 + M_1) : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{H})$ maps the pure state $|\alpha_0\psi_0 + \alpha_1\psi_1\rangle\langle\alpha_0\psi_0 + \alpha_1\psi_1|$ and the mixed state $|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|$ to the same final state. Since unitary maps preserve purity, there can not exist a unitary map $U : \mathcal{H} \to \mathcal{H}$ such that for all $\rho \in \mathcal{T}(\mathcal{H})$:

$$(M_0 + M_1)(\rho) = U\rho U^* .$$

However, according to Schrödinger's equation the development of a closed quantum system is given by a unitary operator. We conclude that the decoherence (2.3) is impossible in a closed system. On the other hand decoherence is a well known and experimentally confirmed phenomenon.

We will therefore consider open systems, i.e. quantum systems which do not obey the Schrödinger equation, but are part of a larger system which does. It has often been pointed out (e.g. [Neu], [Zur]) that decoherence can well occur in this situation, provided that states are only evaluated on the observables of the smaller system. We are more ambitious here: we shall prove that this form of 'local' decoherence is not just a *possible*, but an an *unavoidable* consequence of information-transfer out of the open system.

A. Unitary Information Transfer and Decoherence

We assume that the open system has Hilbert space $\mathcal{H}$, and that its algebra of observables is given by $B(\mathcal{H})$, the bounded operators on $\mathcal{H}$. We may then assume that the larger system has Hilbert space $\mathcal{K} \otimes \mathcal{H}$, since the only way to represent $\mathcal{B}(\mathcal{H})$ on a Hilbert space is in the form $A \mapsto \mathbf{1} \otimes A$ [Tak]. We may think[1] of $\mathcal{B}(\mathcal{K})$ as the observable algebra of some ancillary system in contact with our open quantum system. In this context, $\mathcal{H}$ will be referred to as the 'open system', $\mathcal{K}$ as the 'ancilla' and $\mathcal{K} \otimes \mathcal{H}$ as the 'closed system'.

We couple the system to the ancilla during a finite time interval $[0, t]$. Let $\tau \in \mathcal{T}(\mathcal{K})$ denote the state of the ancilla at time 0, and $\rho \in \mathcal{T}(\mathcal{H})$ that of the small system. The effect of the interaction is described by a unitary operator $U : \mathcal{K} \otimes \mathcal{H} \to \mathcal{K} \otimes \mathcal{H}$, and the state of the pair at time t is given by $U(\tau \otimes \rho)U^* \in \mathcal{T}(\mathcal{K} \otimes \mathcal{H})$. For convenience, we will define the information transfer map $T : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{K} \otimes \mathcal{H})$ by $T(\rho) := U(\tau \otimes \rho)U^*$.

---

[1]Sometimes it may happen, as for instance in fermionic systems, that the observables of the ancilla do not all commute with those of the open system. Also the observable algebra on $\mathcal{K}$ may be smaller than $\mathcal{B}(\mathcal{K})$, but we will neglect these complications here.

In the above setup, we are interested in distinguishing whether the open system $\mathcal{H}$ was in state $|\psi_0\rangle$ or $|\psi_1\rangle$ at time 0. This can be done if there exists a 'pointer observable' $B \otimes \mathbf{1}$ in the ancilla $B(\mathcal{K})$ which takes average value b_0 in state $T(|\psi_0\rangle\langle\psi_0|)$ and b_1 in state $T(|\psi_1\rangle\langle\psi_1|)$. By looking only at the ancilla $\mathcal{K}$ at time t, we are then able to gain information on the state of the open system $\mathcal{H}$ at time 0. We say that information is *transferred* from $\mathcal{H}$ to $\mathcal{K}$.

Under these circumstances, we wish to prove that decoherence occurs on the open system. We prepare the ground by proving the following lemma.

Lemma 2 . *Let ϑ_0, ϑ_1 be unit vectors in a Hilbert space $\mathcal{L}$, and let A and B be bounded self-adjoint operators on $\mathcal{L}$ satisfying $\|[A, B]\| \leq \delta \|A\| \cdot \|B\|$. For $j = 0$ or 1, let $b_j := \langle\theta_j, B\theta_j\rangle$ denote the expectation and $\sigma_j^2 := \langle\theta_j, B^2\theta_j\rangle - \langle\theta_j, B\theta_j\rangle^2$ the variance of B in the state ϑ_j. Then, if $b_0 \neq b_1$,*

$$\big|\langle\vartheta_0, A\vartheta_1\rangle\big| \leq \frac{\delta\|B\| + \sigma_0 + \sigma_1}{|b_0 - b_1|}\|A\| \,.$$

Proof. Since $\|(B - b_j)\vartheta_j\|^2 = \langle\vartheta_j, (B - b_j)^2\vartheta_j\rangle = \sigma_j^2$, we have, by the Cauchy-Schwarz inequality,

$$\big|(b_0 - b_1)\langle\vartheta_0, A\vartheta_1\rangle\big| = \big|\langle\vartheta_0, \big(A(B - b_1) - (B - b_0)A + [B, A]\big)\vartheta_1\rangle\big| \leq \|A\|(\sigma_1 + \sigma_0) + \delta\|A\| \cdot \|B\| \,.$$

□

Note that, for $\delta = \sigma_0 = \sigma_1 = 0$, Lemma 2 merely states that commuting operators respect each other's eigenspaces. We proceed to prove that information transfer causes decoherence on the open system. (See [Jan].)

Theorem 3 *Let ψ_0 and ψ_1 be mutually orthogonal unit vectors in a Hilbert space $\mathcal{H}$, and let $\tau \in \mathcal{T}(\mathcal{K})$ be a state on a Hilbert space $\mathcal{K}$. Let $U : \mathcal{K} \otimes \mathcal{H} \to \mathcal{K} \otimes \mathcal{H}$ be unitary and define $T : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{K} \otimes \mathcal{H})$ by $T(\rho) = U(\tau \otimes \rho)U^*$. Let B be a bounded self-adjoint operator on $\mathcal{K} \otimes \mathcal{H}$, and denote by b_j and σ_j^2 its expected value and variance in the state $T(|\psi_j\rangle\langle\psi_j|)$ for $j = 0, 1$. Suppose that $b_0 \neq b_1$. Then for all $\psi = \alpha_0\psi_0 + \alpha_1\psi_1$ with $|\alpha_0|^2 + |\alpha_1|^2 = 1$ and for all bounded self-adjoint operators A on $\mathcal{K} \otimes \mathcal{H}$ such that $\|[A, B]\| \leq \delta\|A\| \cdot \|B\|$, we have*

$$\Big|\mathrm{tr}\,\big(T(|\psi\rangle\langle\psi|)A\big) - \mathrm{tr}\,\Big(T\big(|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|\big)A\Big)\Big| \leq \frac{\delta\|B\| + \sigma_0 + \sigma_1}{|b_0 - b_1|}\|A\| \,. \tag{3.1}$$

Proof. First, we prove 3.1 in the special case that $\tau = |\varphi\rangle\langle\varphi|$ for some vector $\varphi \in \mathcal{K}$. We introduce the notation $\theta_j := U(\varphi \otimes \psi_j)$. Recall that the expectation of B is given by $b_j = \mathrm{tr}\,\big(T(|\psi_j\rangle\langle\psi_j|)B\big)$, and its variance by $\sigma_j^2 = \mathrm{tr}\,\big(T(|\psi_j\rangle\langle\psi_j|)B^2\big) - \mathrm{tr}^{\,2}\big(T(|\psi_j\rangle\langle\psi_j|)B\big)$. In terms of θ_j, this reduces to $b_j = \langle\theta_j, B\theta_j\rangle$ and $\sigma_j^2 = \langle\theta_j, B^2\theta_j\rangle - \langle\theta_j, B\theta_j\rangle^2$. Similarly, the l.h.s. of 3.1 equals $|\overline{\alpha_0}\alpha_1\langle\theta_0, A\theta_1\rangle + \alpha_0\overline{\alpha_1}\langle\theta_1, A\theta_0\rangle|$, a quantity bounded by $|\langle\theta_0, A\theta_1\rangle|$ since $2|\alpha_0| \cdot |\alpha_1| \leq 1$. Formula 3.1 is then a direct application of Lemma 2.

To reduce the general case to the case above, we note that a non-pure state τ can always be represented as a vector state. Explicitly, suppose that τ decomposes as $\tau = \sum_{i\in\mathbb{N}} |\beta_i|^2 |\varphi_i\rangle\langle\varphi_i|$. Then define the Hilbert space $\tilde{\mathcal{K}} := \bigoplus_{i\in\mathbb{N}} \mathcal{K}_i$, where each $\mathcal{K}_i$ is a copy of $\mathcal{K}$. Now since $(\bigoplus_{i\in\mathbb{N}} \mathcal{K}_i) \otimes \mathcal{H} \cong \bigoplus_{i\in\mathbb{N}} (\mathcal{K}_i \otimes \mathcal{H})$, we may define, for each $X \in \mathcal{B}(\mathcal{K} \otimes \mathcal{H})$, the operator $\tilde{X} \in \mathcal{B}(\tilde{\mathcal{K}} \otimes \mathcal{H})$ by diagonal action on the components of the sum, i.e. $\tilde{X}\big(\bigoplus_{i\in\mathbb{N}}(k_i \otimes h_i)\big) := \bigoplus_{i\in\mathbb{N}} X(k_i \otimes h_i)$. If we now define the vector $\tilde{\varphi} \in \tilde{\mathcal{K}}$ by $\tilde{\varphi} = \bigoplus_i \beta_i\varphi_i$, then we have for all $X \in \mathcal{K} \otimes \mathcal{H}$ and $\chi \in \mathcal{H}$:

$$\begin{aligned}
\mathrm{tr}\,\big(\tilde{U}(|\tilde{\varphi}\rangle\langle\tilde{\varphi}| \otimes |\chi\rangle\langle\chi|)\tilde{U}^*\tilde{X}\big) &= \langle \textstyle\bigoplus_{i\in\mathbb{N}}(\beta_i\varphi_i \otimes \chi), \tilde{U}^*\tilde{X}\tilde{U}\textstyle\bigoplus_{j\in\mathbb{N}}(\beta_j\varphi_j \otimes \chi)\rangle_{\tilde{\mathcal{K}}\otimes\mathcal{H}} \\
&= \langle \textstyle\bigoplus_{i\in\mathbb{N}}(\beta_i\varphi_i \otimes \chi), \textstyle\bigoplus_{j\in\mathbb{N}} U^*XU(\beta_j\varphi_j \otimes \chi)\rangle_{\tilde{\mathcal{K}}\otimes\mathcal{H}} \\
&= \sum_{i\in\mathbb{N}} |\beta_i|^2 \langle(\varphi_i \otimes \chi), U^*XU(\varphi_i \otimes \chi)\rangle_{\mathcal{K}\otimes\mathcal{H}} \\
&= \sum_{i\in\mathbb{N}} |\beta_i|^2 \mathrm{tr}\,\big(U(|\varphi_i\rangle\langle\varphi_i| \otimes |\chi\rangle\langle\chi|)U^*X\big) \\
&= \mathrm{tr}\,\big(U(\tau \otimes |\chi\rangle\langle\chi|)U^*X\big)
\end{aligned}$$

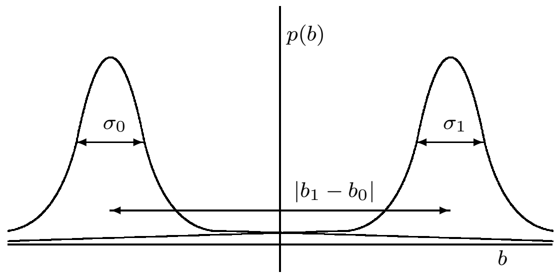
The second step is due to the diagonal action of the operators on $\tilde{\mathcal{K}} \otimes \mathcal{H}$. The problem is now reduced to the vector-case by applying the above to $\chi = \psi$, $\chi = \psi_0$ or $\chi = \psi_1$ and on the other hand $X = A$, $X = B$ or $X = B^2$. □

The backbone of Theorem 3 is formed by the special case $\sigma_0 = \sigma_1 = 0$, $[A, B] = 0$ and $\tau = |\phi\rangle\langle\phi|$, which allows for a short and transparent proof.

In order to arrive at a physical interpretation of Theorem 3, we focus on the case $B = \tilde{B} \otimes \mathbf{1}$, when information is transferred from $\mathcal{H}$ to $\mathcal{K}$. Indeed, examining $\mathcal{K}$ at time t yields information about $\mathcal{H}$ at time 0.

2. Quality of Information Transfer

A small ratio $\frac{\sigma_0+\sigma_1}{|b_0-b_1|}$ indicates a good quality of information transfer. The ratio equals 0 in the perfect case, when $\sigma_0 = \sigma_1 = 0$. Thus $\tilde{B} \otimes \mathbf{1}$ takes a definite value of either b_0 or b_1, depending on whether the initial state of $\mathcal{H}$ was $|\psi_0\rangle$ or $|\psi_1\rangle$. In this case, one can infer the initial state of $\mathcal{H}$ with certainty by inspecting only the ancilla $\mathcal{K}$. More generally, it is still possible to reliably determine from the ancilla $\mathcal{K}$ whether the open system $\mathcal{H}$ was initially in state $|\psi_0\rangle$ or $|\psi_1\rangle$ as long as the standard deviations are small compared to the difference in mean, $\sigma_0, \sigma_1 \ll |b_0 - b_1|$.



Probability densities p of B according to input $|\psi_0\rangle$ and $|\psi_1\rangle$

As the ratio increases, the restriction 3.1 gets less severe, reaching triviality at $\sigma_0 + \sigma_1 = 2|b_0 - b_1|$.

3. Decoherence on the Commutant of the Pointer

Assume perfect information transfer, i.e. $\sigma_0 = \sigma_1 = 0$. If $[A, B] = 0$, then Theorem 3 says that coherent and mixed initial states yield identical distributions of A at time t. In order to distinguish, at time t, whether or not $\mathcal{H}$ was in a pure state at time 0, we will have to use observables A which do not commute with B. But then A and B cannot be observed simultaneously. Summarizing:

> *At time t, it is possible to distinguish whether $\mathcal{H}$ was in state ψ_0 or ψ_1 at time 0. It is also possible to distinguish whether $\mathcal{H}$ was in state ψ or $|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|$ at time 0. But it is not possible to do both.*

We emphasize that this holds even when one has all observables of the entire closed system $\mathcal{K} \otimes \mathcal{H}$ at one's disposal.

4. Decoherence on the Open System

We consider the final state of the open system $\mathcal{H}$, obtained from the final state of the closed system $\mathcal{K} \otimes \mathcal{H}$ by tracing out the degrees of freedom of the ancilla $\mathcal{K}$: an initial state $\rho \in S(\mathcal{H})$ yields final state $\mathrm{tr}_{\mathcal{K}}(T(\rho)) \in S(\mathcal{H})$.

Suppose that information is transferred to a pointer $B = \tilde{B} \otimes \mathbf{1}$ in the ancilla $\mathcal{K}$ with perfect quality, $\sigma_0 = \sigma_1 = 0$. Since $[\mathbf{1} \otimes \tilde{A}, \tilde{B} \otimes \mathbf{1}] = 0$, we see from Theorem 3 that we have $\mathrm{tr}\,(T(|\psi\rangle\langle\psi|)(\mathbf{1} \otimes \tilde{A})) = \mathrm{tr}\,(T(|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|)(\mathbf{1} \otimes \tilde{A}))$ for *all* $\tilde{A} \in \mathcal{B}(\mathcal{H})$, or equivalently

$$\mathrm{tr}_{\mathcal{K}}\big(T(|\psi\rangle\langle\psi|)\big) = \mathrm{tr}_{\mathcal{K}}\Big(T\big(|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|\big)\Big) \,. \tag{3.2}$$

In words:

> *Suppose that at time t, by making a hypothetical measurement of $\tilde{B}$ on the ancilla, it would be possible to distinguish perfectly whether the open system had been in state ψ_0 or ψ_1 at time 0. Then, by looking only at the observables of the open system, it is not possible to distinguish whether $\mathcal{H}$ had been in the pure state $\psi = \alpha_0\psi_0 + \alpha_1\psi_1$ or the collapsed state $|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|$ at time 0.*

This statement holds true, regardless whether $\tilde{B}$ is actually measured or not. (So we do not assume here that such a measurement is physically possible.) We have shown that the map $M_0 + M_1 = \mathrm{tr}_{\mathcal{K}} \circ T$, with $T : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{K} \otimes \mathcal{H})$ the information-transfer operation defined by $T(\rho) := U(\tau \otimes \rho)U^*$, constitutes a physical realization of the abstract decoherence mapping $(M_0 + M_1)$ of section II.

All in all, we have proven that decoherence is an unavoidable consequence of information transfer out of an open system.

5. *Example*

The simplest possible example of unitary information transfer is the following. Let $\mathcal{K} \sim \mathcal{H} \sim \mathbb{C}^2$ be the Hilbert space of a qubit; let $\psi_0 = (1, 0)$ and $\psi_1 = (0, 1)$ be the 'computational basis', and let $U : \mathbb{C}^2 \otimes \mathbb{C}^2 \to \mathbb{C}^2 \otimes \mathbb{C}^2$ be the 'controlled-not gate'. Explicitly, U is defined by $U|\psi_1 \otimes \psi_1\rangle = |\psi_0 \otimes \psi_1\rangle$, $U|\psi_0 \otimes \psi_1\rangle = |\psi_1 \otimes \psi_1\rangle$, $U|\psi_1 \otimes \psi_0\rangle = |\psi_1 \otimes \psi_0\rangle$, and $U|\psi_0 \otimes \psi_0\rangle = |\psi_0 \otimes \psi_0\rangle$. That is, it flips the first qubit whenever the second qubit is set to 1. Let τ be the 0 state of the first qubit.

Since the initial state of the second qubit can be read off from the first, this situation satisfies the hypotheses of Theorem 3 with $B = \sigma_z \otimes \mathbf{1}$ and $\sigma_0 = \sigma_1 = 0$. We verify equation 3.2. For any state $|\psi\rangle = \alpha_0|\psi_0\rangle + \alpha_1|\psi_1\rangle$:

$$\begin{aligned} U|\psi_0 \otimes \psi\rangle &= \alpha_0|\psi_0 \otimes \psi_0\rangle + \alpha_1|\psi_1 \otimes \psi_1\rangle \;:=\; |\theta\rangle; \\ \mathrm{tr}_{\mathcal{K}}\,(|\theta\rangle\langle\theta|) &= |\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|\,. \end{aligned}$$

Thus we have $\mathrm{tr}_{\mathcal{K}}\big(T(|\psi\rangle\langle\psi|)\big) = |\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|$. This agrees with equation 3.2, since one can easily check that $\mathrm{tr}_{\mathcal{K}}\big(T(|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|)\big)$ equals $|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|$ as well.

B. Unitary Information Transfer and State Collapse

We have derived that, in the context of information transfer to an ancillary system, the initial states $|\psi\rangle\langle\psi|$ and $|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|$ lead to the same final state. This is decoherence.

State collapse is a much stronger statement: if outcome '0' is observed, then the system will further behave as if its initial state had been ψ_0 instead of ψ. Similarly, if outcome '1' is observed, then the system will behave as if its initial state had been ψ_1. Now suppose that we ignore the outcome. Since '0' happens with probability $|\alpha_0|^2$ and '1' with probability $|\alpha_1|^2$, the system will behave as if its initial state had been $|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|$. We see that collapse implies decoherence.

The converse does not hold however: imagine a Stern-Gerlach experiment, in which a beam of particles in a σ_x-eigenstate is split in two according to spin in the z-direction. State collapse is the statement that one beam consists of particles with positive spin, the other of particles with negative spin and that both beams have equal intensity. Decoherence is the statement that both outgoing beams *together* consist for 50% of positive-spin particles and for 50% of negative-spin particles. The former statement is strictly stronger than the latter, and deserves separate investigation.

We will therefore answer the following question: suppose that we transfer information to an ancilla $\mathcal{K}$, and then separate $\mathcal{K}$ from $\mathcal{H}$, dividing $\mathcal{H}$ into subensembles according to outcome. What states do we use to describe these subensembles?

1. *Joint Probability Distributions*

A special case of an observable is an *event* p, which in quantum mechanics is represented by a projection P. The relative frequency of occurrence of p is given by $\mathbb{P}(p = 1) = \mathrm{tr}\,(\rho P)$.

The projection $\mathbf{1} - P$ is interpreted as 'not p'. Furthermore, if a projection Q corresponding to an observable q commutes with P, then PQ is again a projection. According to quantum mechanics, p and q can then be observed simultaneously, and the projection PQ is interpreted as the event 'p and q are both observed'.

A state ρ therefore induces a joint probability distribution on p and q:

$$\begin{aligned} \mathrm{tr}\,(\rho PQ) &= \mathbb{P}(p = 1, q = 1)\ ,\ \mathbb{P}(p = 0, q = 1) = \mathrm{tr}\,(\rho(\mathbf{1} - P)Q) \\ \mathrm{tr}\,(\rho P(\mathbf{1} - Q)) &= \mathbb{P}(p = 1, q = 0)\ ,\ \mathbb{P}(p = 0, q = 0) = \mathrm{tr}\,(\rho(\mathbf{1} - P)(\mathbf{1} - Q)) \end{aligned}$$

Particularly relevant is the case in which ρ is a state on a combined space $\mathcal{K}\otimes\mathcal{H}$, and the projections are of the form $Q\otimes\mathbf{1}$ and $\mathbf{1}\otimes P$. (The commuting projections are properties of different systems.) We then have $\mathbb{P}(p=1,q=1) = \mathrm{tr}\,((\mathbf{1}\otimes P)(Q\otimes\mathbf{1})\rho) = \mathrm{tr}\,(P\,\mathrm{tr}_{\mathcal{K}}((Q\otimes\mathbf{1})\rho))$. This holds for all projections P on $\mathcal{H}$, so that the normalized version of $\mathrm{tr}_{\mathcal{K}}((Q\otimes\mathbf{1})\rho)\in\mathcal{T}(\mathcal{H})$ must be interpreted as the state of $\mathcal{H}$, given that $q=1$. Similarly, the normalized version of $\mathrm{tr}_{\mathcal{K}}((\mathbf{1}-Q)\otimes\mathbf{1}\rho)\in\mathcal{T}(\mathcal{H})$ is the state of $\mathcal{H}$, given that $q=0$ is observed.

2. Collapse

Let $T:\rho\mapsto U(\tau\otimes\rho)U^*$ from $\mathcal{T}(\mathcal{H})$ to $\mathcal{T}(\mathcal{K}\otimes\mathcal{H})$ be an information transfer from $\mathcal{H}$ to a pointer-projection $Q\in B(\mathcal{K})$. That is, $\mathrm{tr}\,((Q\otimes\mathbf{1})T(|\psi_0\rangle\langle\psi_0|))=0$ and $\mathrm{tr}\,((Q\otimes\mathbf{1})T(|\psi_1\rangle\langle\psi_1|))=1$, so that at time t, one can see from $\mathcal{K}$ whether $\mathcal{H}$ was in state ψ_0 or ψ_1 at time 0.

Since $Q\otimes\mathbf{1}$ commutes with all of $\mathbf{1}\otimes B(\mathcal{H})$, it is possible to separate $\mathcal{H}$ from $\mathcal{K}$, and divide $\mathcal{H}$ into subensembles according to the outcome of Q. This is done as follows: with any measurement on $\mathcal{H}$, a simultaneous measurement of Q on $\mathcal{K}$ is performed to determine in which ensemble this particular system should fall. It follows from the above that the 1-ensemble should be described by the normalized version of $M_1(\rho):=\mathrm{tr}_{\mathcal{K}}((Q\otimes\mathbf{1})T(\rho))$, and the 0-ensemble by the normalized version of $M_0(\rho):=\mathrm{tr}_{\mathcal{K}}((\mathbf{1}-Q\otimes\mathbf{1})T(\rho))$. Since Q commutes with $B(\mathcal{H})$, this is just conditioning on a classical probability space at time t. We have arrived at an interpretation of the map $M(\rho):=M_0(\rho)\oplus M_1(\rho)$ of section II.

We will now prove that M takes the form $M(|\psi\rangle\langle\psi|)=|\alpha_0|^2\mathrm{tr}_{\mathcal{K}}T(|\psi_0\rangle\langle\psi_0|)\oplus|\alpha_1|^2\mathrm{tr}_{\mathcal{K}}T(|\psi_1\rangle\langle\psi_1|)$. This is a strong physical statement. For instance, any spin-system $\alpha_0|\psi_0\rangle+\alpha_1|\psi_1\rangle$ that is found to have spin 1 in the z-direction may subsequently be treated as if it had been in state ψ_1 at time 0. This is nontrivial: a priori, it is perfectly conceivable that the different initial states ψ_0 and ψ result in different final states, even though they yield the same Q-output.

One could alternatively, (and more traditionally), arrive at the 'collapse of the wavefunction' $M(|\psi\rangle\langle\psi|)=|\alpha_0|^2\mathrm{tr}_{\mathcal{K}}T(|\psi_0\rangle\langle\psi_0|)\oplus|\alpha_1|^2\mathrm{tr}_{\mathcal{K}}T(|\psi_1\rangle\langle\psi_1|)$ by assuming that, at time 0, the quantum system makes either the jump $|\psi\rangle\langle\psi|\mapsto|\psi_0\rangle\langle\psi_0|$ or the jump $|\psi\rangle\langle\psi|\mapsto|\psi_1\rangle\langle\psi_1|$. Since we arrive at the same conclusion, namely the above 'collapse of the wavefunction', using only open systems, unitary transformations and the probabilistic interpretation of quantum mechanics, such an assumption of 'jumps' at time 0 is made redundant.

Proposition 4 *Let $T:\rho\mapsto U(\tau\otimes\rho)U^*$ from $\mathcal{T}(\mathcal{H})$ to $\mathcal{T}(\mathcal{K}\otimes\mathcal{H})$ satisfy $\mathrm{tr}\,((Q\otimes\mathbf{1})T(|\psi_0\rangle\langle\psi_0|))=0$ and $\mathrm{tr}\,((Q\otimes\mathbf{1})T(|\psi_1\rangle\langle\psi_1|))=1$ for some 'pointer-projection' Q on $\mathcal{K}$. Define a map $M:\mathcal{T}(\mathcal{H})\mapsto\mathcal{T}(\mathcal{H})\oplus\mathcal{T}(\mathcal{H})$ by $M(\rho):=\mathrm{tr}_{\mathcal{K}}((\mathbf{1}-Q\otimes\mathbf{1})T(\rho))\oplus\mathrm{tr}_{\mathcal{K}}((Q\otimes\mathbf{1})T(\rho))$. Then for $\psi=\alpha_0\psi_0+\alpha_1\psi_1$ we have $M(|\psi\rangle\langle\psi|)=|\alpha_0|^2\mathrm{tr}_{\mathcal{K}}T(|\psi_0\rangle\langle\psi_0|)\oplus|\alpha_1|^2\mathrm{tr}_{\mathcal{K}}T(|\psi_1\rangle\langle\psi_1|)$.*

This can be seen almost directly from Proposition 1:

Proof. Since $M_1(|\psi_0\rangle\langle\psi_0|)\geq 0$ is a positive operator, we may conclude from $\mathrm{tr}\,(M_1(|\psi_0\rangle\langle\psi_0|))=0$ that $M_1(|\psi_0\rangle\langle\psi_0|)=0$. Similarly $M_0(|\psi_1\rangle\langle\psi_1|)=0$. Utilizing Proposition 1, we find that $M(|\psi\rangle\langle\psi|)=|\alpha_0|^2M_0(|\psi_0\rangle\langle\psi_0|)\oplus|\alpha_1|^2M_1(|\psi_1\rangle\langle\psi_1|)$. The proof is completed by noting from $\mathrm{tr}_{\mathcal{K}}((\mathbf{1}-Q\otimes\mathbf{1})T(|\psi_1\rangle\langle\psi_1|))=0$ that $\mathrm{tr}_{\mathcal{K}}(T(|\psi_1\rangle\langle\psi_1|))=\mathrm{tr}_{\mathcal{K}}((Q\otimes\mathbf{1})T(|\psi_1\rangle\langle\psi_1|))=M_1(|\psi_1\rangle\langle\psi_1|)$, and similarly that $\mathrm{tr}_{\mathcal{K}}(T(|\psi_0\rangle\langle\psi_0|))=M_0(|\psi_0\rangle\langle\psi_0|)$. □

We summarize:

> *Consider an ensemble of systems of type $\mathcal{H}$ in state ψ. Suppose that information is transferred to a pointer-projection Q on an ancillary system $\mathcal{K}$. Subsequently, the ensemble is divided into two subensembles according to outcome. Then all observations on $\mathcal{H}$ made afterwards, conditioned on the observation that the measurement outcome was 0, will be as if the system had originally been in the collapsed state ψ_0 instead of ψ. No independent 'collapse postulate' is needed to arrive at this conclusion.*

3. Example

In the simple model of information transfer introduced in Section III A, we will now demonstrate why repeated spin-measurements yield identical outcomes.

The probed system is once again a single spin $\mathcal{H}=\mathbb{C}^2$, whereas the ancillary system now consists of two spins, $\mathcal{K}=\mathbb{C}^2\otimes\mathbb{C}^2$ in initial state $|\psi_0\otimes\psi_0\rangle$. Repeated information-transfer, first to pointer $\sigma_{z,1}$ and then to $\sigma_{z,2}$, is then represented by the unitary $U:=U_2U_1$ on $\mathcal{K}\otimes\mathcal{H}$. In this expression, U_1 is the controlled not-gate flipping the first qubit of $\mathcal{K}$ if $\mathcal{H}$ is set to 1, and U_2 flips the second qubit of $\mathcal{K}$ if $\mathcal{H}$ is set to 1.

Since $U|\psi_0\otimes\psi_0\otimes(\alpha_0\psi_0+\alpha_1\psi_1)\rangle=|\alpha_0\psi_0\otimes\psi_0\otimes\psi_0\rangle+|\alpha_1\psi_1\otimes\psi_1\otimes\psi_1\rangle$, we can explicitly calculate the joint probability distribution on the two pointers $\sigma_{z,1}$ and $\sigma_{z,2}$ in the final state:

$$\begin{array}{lll} \mathbb{P}(s_{z,1} = \ \ \ 1, s_{z,2} = 1) = |\alpha_1|^2 & , \quad 0 & = \mathbb{P}(s_{z,1} = \ \ \ 1, s_{z,2} = -1) \\ \mathbb{P}(s_{z,1} = -1, s_{z,2} = 1) = \ \ 0 & , \ |\alpha_0|^2 & = \mathbb{P}(s_{z,1} = -1, s_{z,2} = -1) \end{array}$$

In particular, we see that if the first outcome is 1 (which happens with probability $|\alpha_1|^2$), then so is the second. Proposition 4 shows that this is the general situation, independent of the (rather simplistic) details of this particular model.

C. Information Leakage to the Environment

On closed systems decoherence does not occur, because unitary time evolution preserves the purity of states. However, macroscopic systems are almost never closed.

Imagine, for example, that $\mathcal{H} = \mathbb{C}^2$ represents a two-level atom, and $\mathcal{K}$ some large measuring device. Information about the energy $\mathbf{1} \otimes \sigma_z$ of the atom is transferred to the apparatus, where it is stored as the position $\tilde{B} \otimes \mathbf{1}$ of a pointer. Then as soon as information on the pointer-position $\tilde{B} \otimes \mathbf{1}$ leaves the system, collapse on the combined atom-apparatus system takes place. For example, a ray of light may reflect on the pointer, revealing its position to the outside world. (See [J&Z].) It is of course immaterial whether or not someone is actually *looking* at the photons. If even the smallest speck of light were to fall on the pointer, the information about the pointer position would already be encoded in the light, causing full collapse on the atom-apparatus system. (See [Zur] for an example.)

The quality of this information transfer will not be perfect. If a macroscopic system is interacting normally with the outside world, (the occasional photon happens to scatter on it, for instance), then a number of macroscopic observables X will leak information continually, with a macroscopic uncertainty σ. This enables us to apply Theorem 3. It says that all coherences between eigenstates ψ_{x_1} and ψ_{x_2} of macroscopic observables X are continually vanishing on the macroscopic system $\mathcal{L}$, provided that their eigenvalues x_1 and x_2 satisfy $|x_1 - x_2| \gg 2\sigma$. (The pointer, e.g. a beam of light, is outside the system, so that $\delta = 0$.)

Take for example a collection of N spins, $\mathcal{L} = \bigotimes_{i=1}^{N} \mathbb{C}^2$. Suppose that for $\alpha = x, y, z$, the average spin-observables $S_\alpha = \frac{1}{N} \sum_{i=1}^{N} \sigma_\alpha^i$ are continually being measured with an accuracy[2] $N^{-\frac{1}{2}} \ll \sigma \ll 1$. Then between macroscopically different eigenstates of S_α, i.e. states for which the eigenvalues satisfy $|s_\alpha - s'_\alpha| \gg \sigma$, coherences are constantly disappearing. However, the information leakage need not have any effect on states which only differ on a microscopic scale. Take for instance $\rho \otimes |+\rangle\langle+|$ and $\rho \otimes |-\rangle\langle-|$, with ρ an arbitrary state on $N-1$ spins. Indeed, $|s_\alpha - s'_\alpha| \leq 2/N \ll \sigma$, so Theorem 3 is vacuous in this case: no decoherence occurs.

We see how the variance σ^2 produces a smooth boundary between the macroscopic and the microscopic world: macroscopically distinguishable states (involving S_α-differences $\gg \sigma$) continually suffer from loss of coherence, while states that only differ microscopically (involving S_α-differences $\ll \sigma$) are unaffected.

In case of a system monitored by a macroscopic measurement apparatus, we are interested in coherence between eigenstates of the macroscopic pointer. By definition, these eigenstates are macroscopically distinguishable. We may then give the following answer to the question why it is so hard, in practice, to witness coherence:

> *If information leaks from the pointer into the outside world, decoherence takes place on the combination of system and measurement apparatus. In practice, macroscopic pointers constantly leak information.*

IV. CLOSED SYSTEMS

Closed systems evolve according to unitary time evolution, so that coherence which is present initially will still be there at later times. Yet on macroscopic systems, coherent superpositions are almost never observed. Why is this the case?

---

[2]Since $[S_x, S_y] \neq 0$, they cannot be simultaneously measured with complete accuracy, see e.g. [Wer]. However, this problem disappears if the accuracy satisfies $\sigma^2 \geq \frac{1}{2} \|[S_x, S_y]\| = \frac{1}{N}$, see [Jan]. For large N, (typically $N \sim 6 \times 10^{23}$), this allows for extremely accurate measurement.

A. Macroscopic Systems

Because of the direct link that it provides between the scale of a system on the one hand, and on the other hand the difficulties in witnessing coherence, we feel that the following line of reasoning, essentially due to Hepp [Hep], is the most important mechanism hiding coherence.

Let us first define what we mean by macroscopic and microscopic observables. We consider a system consisting of N distinct subsystems, i.e. $\mathcal{K} = \bigotimes_{i=1}^{N} \mathcal{K}_i$. If one thinks of $\mathcal{K}_i$ as the atoms out of which a macroscopic system $\mathcal{K}$ is constructed, N may well be in the order of 10^{23}.

We will define the *microscopic* observables to be the ones that refer only to one particular subsystem $\mathcal{K}_i$:

Definition. An observable $X \in B(\mathcal{K})$ is called *microscopic* if it is of the form $X = \mathbf{1} \otimes \ldots \otimes \mathbf{1} \otimes X_i \otimes \mathbf{1} \otimes \ldots \otimes \mathbf{1}$ for some $i \in \{1, 2, \ldots, N\}$ and some $X_i \in B(\mathcal{K}_i)$.

In this situation we will identify $X_i \in B(\mathcal{K}_i)$ with $X \in B(\mathcal{K})$. We take macroscopic observables to be averages of microscopic observables 'of the same size':

Definition. An observable $Y \in B(\mathcal{K})$ is called *macroscopic* if it is of the form $Y = \frac{1}{N}\sum_{i=1}^{N} Y_i$, with $Y_i \in B(\mathcal{K}_i)$ such that $\|Y_i\| \leq \|Y\|$.

We will only use the term 'macroscopic' in this narrow sense from here on, even though there do exist observables which are called 'macroscopic' in daily life, but do not fall under the above definition.

Now suppose that we transfer information from a system $\mathcal{H}$ to a macroscopic system $\mathcal{K} = \bigotimes_{i=1}^{N} \mathcal{K}_i$, using a macroscopic pointer $\tilde{B} \in B(\mathcal{K})$. As explained before, we then have a map $T : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{K} \otimes \mathcal{H})$ such that the pointer $\tilde{B} \otimes \mathbf{1}$ has different expectation values b_0 and b_1 in the states $T(|\psi_0\rangle\langle\psi_0|)$ and $T(|\psi_1\rangle\langle\psi_1|)$.

Since $\tilde{B}$ is macroscopic, it is unrealistic to require $T(|\psi_0\rangle\langle\psi_0|)$ and $T(|\psi_1\rangle\langle\psi_1|)$ to be eigenstates of $\tilde{B}$. Instead, we will require their standard deviations in $\tilde{B}$ to be negligible compared to their difference in mean, i.e. $\sigma_0 \ll |b_0 - b_1|$ and $\sigma_1 \ll |b_0 - b_1|$.

After this information transfer, we try to distinguish whether the system $\mathcal{H}$ had initially been in the coherent state $\alpha_0|\psi_0\rangle + \alpha_1|\psi_1\rangle$ or in the incoherent mixture $|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|$. We have already shown that this cannot be done by measuring observables in $\mathbf{1} \otimes B(\mathcal{H})$. The following adaptation of Theorem 3 shows that it is also impossible to do this by measuring macroscopic or microscopic observables on the closed system $\mathcal{K} \otimes \mathcal{H}$.

Corollary 5 *Let ψ_0 and ψ_1 be orthogonal unit vectors in a Hilbert space $\mathcal{H}$ and let $\tau \in \mathcal{T}(\mathcal{K})$ be a state on the Hilbert space $\mathcal{K} = \bigotimes_{i=1}^{N} \mathcal{K}_i$. Let $U : \mathcal{K} \otimes \mathcal{H} \to \mathcal{K} \otimes \mathcal{H}$ be unitary and define $T : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{K} \otimes \mathcal{H})$ by $T(\rho) = U(\tau \otimes \rho)U^*$. Let $\tilde{B}$ be a macroscopic observable in $B(\mathcal{K})$, and define $B := \tilde{B} \otimes \mathbf{1}$. Denote by b_j and σ_j^2 its expected value and variance in the state $T(|\psi_j\rangle\langle\psi_j|)$ for $j = 0, 1$. Suppose that $b_0 \neq b_1$. Then for all $\psi = \alpha_0\psi_0 + \alpha_1\psi_1$ with $|\alpha_0|^2 + |\alpha_1|^2 = 1$ and for all microscopic and macroscopic observables $A \in B(\mathcal{K} \otimes \mathcal{H})$, we have*

$$\Big|\mathrm{tr}\left(T(|\psi\rangle\langle\psi|)A\right) - \mathrm{tr}\left(T\left(|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1^2||\psi_1\rangle\langle\psi_1|\right)A\right)\Big| \leq \frac{\frac{2}{N}\|B\| + \sigma_0 + \sigma_1}{|b_0 - b_1|}\|A\|\,.$$

Proof. If A is microscopic, we have $\|[A, B]\| = \|[A_i, \frac{1}{N}\sum_{j=1}^{N} B_j]\| = \frac{1}{N}\|[A_i, B_i]\| \leq \frac{2\|A\|\|B\|}{N}$. If A is macroscopic, we have $\|[A, B]\| = \|[\frac{1}{N+1}\sum_{i=0}^{N} A_i, \frac{1}{N}\sum_{j=1}^{N} B_j]\| = \frac{1}{N(N+1)}\sum_{i=1}^{N}\|[A_i, B_i]\| \leq \frac{2\|A\|\|B\|}{N}$. Either way, we can now apply Theorem 3. □

B. Examples

In order to illustrate the above, we discuss four examples of information transfer to a macroscopic system.

1. The Finite Spin-Chain

We study a single spin $\mathcal{H} = \mathbb{C}^2$ in interaction with a large but finite spin-chain $\mathcal{K} = \bigotimes_{i=1}^{N} \mathbb{C}^2$, the latter acting as a measurement apparatus. Once again, let $\psi_0 = (1, 0)$ and $\psi_1 = (0, 1)$ be the 'computational basis'. Initially, all spins in the spin-chain are down: $\tau = |\psi_0 \otimes \ldots \otimes \psi_0\rangle\langle\psi_0 \otimes \ldots \otimes \psi_0|$. Let $U_i : \mathcal{K} \otimes \mathcal{H} \to \mathcal{K} \otimes \mathcal{H}$ be the 'controlled-not gate', which flips spin number i in the chain whenever the single qubit is set to 1. (We define $U_j = \mathbf{1}$ for $j \notin \{1, 2, \ldots, N\}$.)

$$U_i = \mathbf{1} \otimes P_- + \sigma_{x,i} \otimes P_+ \quad \text{with} \quad P_+ = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad P_- = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

In discrete time $n \in \mathbb{Z}$, the unitary evolution is given by $n \mapsto U_n U_{n-1} \ldots U_2 U_1$. (See [Hep].) This represents a single spin flying over a spin-chain from 1 to N, interacting with spin n at time n.

Obviously $U_N|\psi_0 \otimes \ldots \otimes \psi_0\rangle \otimes |\psi_0\rangle = |\psi_0 \otimes \ldots \otimes \psi_0\rangle \otimes |\psi_0\rangle$ and $U_N|\psi_0 \otimes \ldots \otimes \psi_0\rangle \otimes |\psi_1\rangle = |\psi_1 \otimes \ldots \otimes \psi_1\rangle \otimes |\psi_1\rangle$. We consider the average spin of the spin-chain as pointer, $B = \frac{1}{N}\sum_{i=1}^{N} \sigma_{z,i}$. This makes the map $T : \rho \mapsto U_N \tau \otimes \rho U_N^*$ an information transfer to a macroscopic system. Applying Corollary 5 with $b_0 = -1$, $b_1 = 1$ and $\sigma_0 = \sigma_1 = 0$ yields the estimate

$$\Big|\mathrm{tr}\,\big(T(|\alpha_0\psi_0 + \alpha_1\psi_1\rangle\langle\alpha_0\psi_0 + \alpha_1\psi_1|)A\big) - \mathrm{tr}\,\Big(T\big(|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1^2||\psi_1\rangle\langle\psi_1|\big)A\Big)\Big| \leq \frac{1}{N}\|A\|$$

for all microscopic and macroscopic $A \in B(\mathcal{K} \otimes \mathcal{H})$. Indeed, in this particular model, the estimated quantity is identically zero since $\langle\psi_0 \otimes \ldots \psi_0, X_i\,\psi_1 \otimes \ldots \otimes \psi_1\rangle = \langle\psi_0, \psi_1\rangle^{N-1}\langle\psi_0, X_i\psi_1\rangle = 0$ for all microscopic X_i.

Of course coherence *can* be detected on the closed system $\mathcal{K} \otimes \mathcal{H}$, but only using observables that are neither macroscopic nor microscopic, such as $\sigma_x \otimes \ldots \otimes \sigma_x$.

2. Finite Spin-Chain at Nonzero Temperature

A more realistic initial state for the spin-chain is the nonzero-temperature state $\tau_\beta = \frac{e^{-\beta H}}{\mathrm{tr}\,e^{-\beta H}}$. For the spin-chain Hamiltonian we will take $H = \sum_i \sigma_{z,i} = NB$, so that τ_β becomes the tensor product of N copies of the $\mathbb{C}^2$-state

$$\hat{\tau}_\beta = \frac{1}{e^{\beta} + e^{-\beta}} \begin{pmatrix} e^{-\beta} & 0 \\ 0 & e^{\beta} \end{pmatrix}.$$

With the same time-evolution as before, we have $T(|\psi_0\rangle\langle\psi_0|) = |\psi_0\rangle\langle\psi_0| \otimes \rho_\beta$ and $T(|\psi_1\rangle\langle\psi_1|) = |\psi_1\rangle\langle\psi_1| \otimes \rho_{-\beta}$. Again we choose the mean energy B as our pointer. A brief calculation shows that $\mathrm{tr}\,(B\tau_\beta) = \frac{e^{-\beta} - e^{\beta}}{e^{\beta} + e^{-\beta}} =: \varepsilon(\beta)$ and that $\mathrm{tr}\,(B^2\tau_\beta) - \mathrm{tr}\,(B\rho_\beta)^2 = \frac{1}{N}(1 - \varepsilon^2(\beta))$. Corollary 5 now gives us, for microscopic and macroscopic A,

$$\Big|\mathrm{tr}\,\big(T(|\alpha_0\psi_0 + \alpha_1\psi_1\rangle\langle\alpha_0\psi_0 + \alpha_1\psi_1|)A\big) - \mathrm{tr}\,\Big(T\big(|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1^2||\psi_1\rangle\langle\psi_1|\big)A\Big)\Big| \leq \left(\frac{1}{\varepsilon(\beta)N} + \frac{\sqrt{1 - \varepsilon^2(\beta)}}{\varepsilon(\beta)\sqrt{N}}\right)\|A\|.$$

For large N, we see that the term $\sim \frac{1}{N}$ due to the fact that $[A, B] \neq 0$ is dominated by the thermodynamical fluctuations, which of course go as $\sim \frac{1}{\sqrt{N}}$. In statistical physics, it is standard practice to neglect even the latter.

3. Energy as a Pointer

Hamiltonians often fail to be macroscopic in our narrow sense of the word, since they are generically unbounded and contain interaction terms. However, this does not imply failure of our scheme to estimate coherence.

For example, consider an N-particle system with Hilbert space $\mathcal{K} = \bigotimes_{i=1}^{N} \mathcal{K}_i$ and Hamiltonian $H = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + V(x_1, x_2, \ldots, x_N)$. Information is transferred from $\mathcal{H}$ to $\mathcal{K}$ with H as pointer, that is the two states $\mathrm{tr}_{\mathcal{H}}(T(|\psi_0\rangle\langle\psi_0|))$ and $\mathrm{tr}_{\mathcal{H}}(T(|\psi_1\rangle\langle\psi_1|))$ have different energies E and E'. Without loss of generality, assume that they are vectorstates: $\mathrm{tr}_{\mathcal{H}}(T(|\psi_0\rangle\langle\psi_0|)) = |\psi\rangle\langle\psi|$ and $\mathrm{tr}_{\mathcal{H}}(T(|\psi_1\rangle\langle\psi_1|)) = |\psi'\rangle\langle\psi'|$. (Density matrices can always be represented as vectors on a different Hilbert space, cf. the proof of Theorem 3.)

We thus have two vector states $|\psi\rangle$ and $|\psi'\rangle$ with different energies $E := \langle\psi, H\psi\rangle$ and $E' := \langle\psi', H\psi'\rangle$. We estimate the coherence between $|\psi\rangle$ and $|\psi'\rangle$ on x_n, the position of particle n.

$$
\begin{aligned}
(E-E')\langle\psi, x_n\psi'\rangle &= \langle E\psi, x_n\psi'\rangle - \langle x_n\psi, E'\psi'\rangle \\
&= \langle H\psi-(H-E)\psi, x_n\psi'\rangle - \langle x_n\psi, H\psi'-(H-E')\psi'\rangle \\
&= \langle [H,x_n]\psi, \psi'\rangle - \langle (H-E)\psi, x_n\psi'\rangle + \langle x_n\psi, (H-E')\psi'\rangle
\end{aligned}
$$

Now since $[H, x_n] = \frac{1}{2m_n}[p_n^2, x_n] = \frac{-i\hbar p_n}{m_n}$, we can apply the Cauchy-Schwarz inequality in each term to obtain

$$
|E-E'|\,|\langle\psi, x_n\psi'\rangle| \leq \frac{\hbar}{m_n}\sqrt{\langle\psi, p_n^2\psi\rangle} + \sqrt{\langle\psi, x_n^2\psi\rangle}\sqrt{\langle\psi', (H-E')^2\psi'\rangle} + \sqrt{\langle\psi', x_n^2\psi'\rangle}\sqrt{\langle\psi, (H-E)^2\psi\rangle}\ .
$$

If we define the characteristic speed $V_n := \sqrt{\langle\psi, (\frac{p_n}{m_n})^2\psi\rangle}$, the characteristic positions $X_n := \sqrt{\langle\psi, x_n^2\psi\rangle}$ and $X_n' := \sqrt{\langle\psi', x_n^2\psi'\rangle}$, and the standard deviations $\sigma := \sqrt{\langle\psi, (H-E)^2\psi\rangle}$ and $\sigma' := \sqrt{\langle\psi', (H-E')^2\psi'\rangle}$, we obtain

$$
|\langle\psi, x_n\psi'\rangle| \leq \frac{\hbar V_n + \sigma X_n' + \sigma' X_n}{|E-E'|}\ .
$$

As such, this doesn't tell us very much. We will have to make some physically plausible assumptions on the state of the system in order to obtain results. First, we assume that the system is encased in an $L \times L \times L$ box so that $X_n, X_n' \leq L$. Also, we assume $V_n < c$. This yields $|\langle\psi, x_n\psi'\rangle| \leq \frac{\hbar c + L(\sigma+\sigma')}{|E-E'|}$. Secondly, we assume that scaling the system in any meaningful way will produce $|E-E'| \sim N$ and $\sigma + \sigma' \sim \sqrt{N}$, so that the coherence on x_n approaches zero as $\sim \frac{1}{\sqrt{N}}$. Notice the almost thermodynamic lack of detail required for this estimate.

4. Schrödinger's Cat

Let us finally analyze the rather drastic extraction of information from a radioactive particle that has become known[3] as 'Schrödinger's cat'. (See [Sch].) The experiment is performed as follows. We are interested in a radioactive particle. Is it in a decayed state ψ_0 or in a non-decayed state ψ_1?

In order to determine this, we set up the following experiment. A Geiger counter is placed next to the radioactive particle. If the particle decays, then the Geiger counter clicks. A mechanism then releases a hammer, which smashes a vial of hydrocyanic acid, killing a cat. All of this happens in a closed box no higher than $1m$, and completely impenetrable to information. A measurement of the atom is done as follows: first, place it inside the box. Then wait for a period of time that is long compared to the decay time of the atom. Finally, open the box, and inspect whether the cat has dropped dead or is still standing upright.

The atom is described by a Hilbert space $\mathcal{H}$, the combination of Geiger counter, mechanism, hammer, vial and cat by a Hilbert space $\mathcal{K}$. Initially, the latter is prepared in a state $|\theta\rangle$. As a pointer, we take the center of mass of the cat, $Z := \frac{1}{N}\sum_{i=1}^{N} z_i$. In this expression, N is the amount of atoms out of which the cat is constructed, and z_i is the z-component of particle number i. (It is a harmless assumption that all atoms in the cat have the same mass.) Since the box only measures $1m$ in height, we may take $\|Z\| = 1$. The unitary evolution $U \in B(\mathcal{K} \otimes \mathcal{H})$ then produces $U|\psi_0 \otimes \theta\rangle := |\gamma_0\rangle$ and $U|\psi_1 \otimes \theta\rangle := |\gamma_1\rangle$, which are eigenstates[4] of Z with different eigenvalues.

Suppose that, initially, the atom is either in the decayed state ψ_0 with probability $|\alpha_0|^2$ or in the non-decayed state ψ_1 with probability $|\alpha_1|^2$. That is, the initial state is the incoherent mixture $|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|$. By linearity, the final state is then the incoherent state $|\alpha_0|^2|\gamma_0\rangle\langle\gamma_0| + |\alpha_1|^2|\gamma_1\rangle\langle\gamma_1|$.

On the other hand, if the atom starts out in the coherent superposition $\alpha_0|\psi_0\rangle + \alpha_1|\psi_1\rangle$, then the combined system ends up in the coherent state $U|(\alpha_0\psi_0 + \alpha_1\psi_1) \otimes \theta\rangle = \alpha_0|\gamma_0\rangle + \alpha_1|\gamma_1\rangle$.

The question is now this: why do we not notice the difference between these two situations if we open the box? First of all, according to Theorem 3 (and the observations following it in section III A 3), it is impossible to detect coherence between γ_0 and γ_1 *and* ascertain the position of the cat. Upon opening the black box, we must make a choice.

---

[3]Actually, Schrödinger's proposal was slightly different. In the original thought experiment, death of the cat was correlated with decay of the atom at time t instead of 0, which wouldn't make it an information transfer in our sense of the word.

[4]As discussed before, it would be more realistic to allow for a nonzero variance $0 < \sigma_j \ll 1$ instead of requiring θ_j to be eigenstates of Z. We use $\sigma_j = 0$ for clarity, leaving the argument essentially unchanged .

Secondly, according to the discussion in section III C, the coherences between the macroscopically different states γ_0 and γ_1 are extremely volatile. Any speck of light falling on the cat will reveal its position with reasonable accuracy, causing the coherence to disappear according to Theorem 3.

Yet even if we were able to open the box without any information on the position of the cat leaking out, even then would we be unable to detect coherence between γ_0 and γ_1. Apply Corollary 5 to the transfer of information from atom to cat. We have $\sigma_0 = \sigma_1 = 0$, and with pointer Z we have $\|Z\| = 1$ (the height of the box is 1 m) and $z_1 - z_0 = 0.1$ (the difference between a cat that is standing up and one that has dropped dead is 10 cm). We then obtain for all macroscopic and microscopic A:

$$\left|\langle \alpha_0\gamma_0 + \alpha_1\gamma_1, A\alpha_0\gamma_0 + \alpha_1\gamma_1\rangle - \left(|\alpha_0|^2\langle\gamma_0, A\gamma_0\rangle + |\alpha_1|^2\langle\gamma_1, A\gamma_1\rangle\right)\right| \leq \frac{20}{N}\|A\| \,.$$

On the subset of observables we are normally able to measure, the distinction between coherent and incoherent mixtures practically vanishes for $N \sim 10^{23}$. For all practical intents and purposes, it is completely harmless to assume that the final state of the cat is $|\alpha_0|^2|\gamma_0\rangle\langle\gamma_0| + |\alpha_1|^2|\gamma_1\rangle\langle\gamma_1|$ instead of $\alpha_0|\gamma_0\rangle + \alpha_1|\gamma_1\rangle$. But it would be false to state that the former has actually been observed.

V. CONCLUSION

In open systems, we have proven that decoherence is a necessary consequence of information transfer to the outside. More in detail, we have reached the following conclusions:

- Suppose that an open system $\mathcal{H}$ interacts with an ancillary system $\mathcal{K}$ in such a way, that it is possible, in principle, to determine from $\mathcal{K}$ whether $\mathcal{H}$ had been in state ψ_0 or ψ_1 before the interaction. If $\mathcal{H}$ started out in a coherent state $\alpha_0|\psi_0\rangle + \alpha_1|\psi_1\rangle$, then it will behave after the information transfer as if it had started out in the incoherent mixture $|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|$ instead. This is called 'decoherence'.

- Suppose again that the information whether $\mathcal{H}$ was in state ψ_0 or ψ_1 is transported to an ancillary system $\mathcal{K}$. This is done with an ensemble of $\mathcal{H}$-systems described by the state $\alpha_0|\psi_0\rangle + \alpha_1|\psi_1\rangle$. The ensemble is then split into subensembles, according to outcome. The '0-ensemble' then behaves as if it had been in state ψ_0 at the beginning of the procedure, and the '1-ensemble' as if it had started in state ψ_1. This is called 'state collapse'.

- These results were obtained entirely within the framework of traditional quantum mechanics and unitary time evolution on a larger, closed system containing $\mathcal{H}$. No 'reduction-postulate' is needed. From Proposition 1, we see that any information extraction causes collapse, quite independent of its particular mechanism.

- On the closed system containing the smaller, open one no decoherence occurs in principle. In practice however, closed systems are very hard to achieve. We have argued that information transfer from a macroscopic observable A, performed with macroscopic precision σ, causes decoherence between eigenstates of A if their values satisfy $\sigma \ll |a_1 - a_0|$. Since information on macroscopic observables tends to leak out, coherence between macroscopically different states tends to vanish.

Still, even if the combined system $\mathcal{K} \otimes \mathcal{H}$ is considered perfectly closed, there are some results to be obtained. Again, we investigated the case that a system $\mathcal{H}$ interacts unitarily with a system $\mathcal{K}$ in such a way that the information whether $\mathcal{H}$ was in state ψ_0 or ψ_1 can be read off from a pointer in $\mathcal{K}$. We have reached the following conclusions concerning the closed system $\mathcal{K} \otimes \mathcal{H}$:

- Using only observables on the closed system that commute with the pointer, it is impossible to detect whether $\mathcal{H}$ had started out in state $\alpha_0|\psi_0\rangle + \alpha_1|\psi_1\rangle$ or $|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|$. Physically, this means that it is impossible to distinguish between coherent and incoherent initial states while at the same time distinguishing between ψ_0 and ψ_1.

- Suppose that the closed system $\mathcal{K}\otimes\mathcal{H}$ is macroscopic, and that one has access to its macroscopic and microscopic observables only. Then it is almost impossible to distinguish whether $\mathcal{H}$ had started out in state $\alpha_0|\psi_0\rangle + \alpha_1|\psi_1\rangle$ or $|\alpha_0|^2|\psi_0\rangle\langle\psi_0| + |\alpha_1|^2|\psi_1\rangle\langle\psi_1|$. We have obtained upper bounds on the coherences $\langle\psi_0, A\psi_1\rangle$, evaluated on microscopic or macroscopic A. Assuming perfect information transfer ($\sigma_0 = \sigma_1 = 0$), they approach zero as $\sim \frac{1}{N}$, where N is the size of the system.

In short: no decoherence ever occurs on perfectly closed systems, even if they are macroscopic. It is just very hard to distinguish coherent from incoherent states, creating the false impression that it does.

The link between decoherence and macroscopic systems was brought forward by Klaus Hepp in his fundamental paper [Hep], where he considered infinite closed systems, displaying decoherence in infinite time. In infinite systems, the microscopic observables form a non-commutative C*-algebra $\mathcal{A}$. Its weak closure $\mathcal{A}''$ is considered as the (von Neumann-)algebra of all observables. The macroscopic observables form a commutative algebra $\mathcal{C}$ which is contained in the centre of $\mathcal{A}''$, i.e. $\mathcal{C} \subset \mathcal{Z} = \{Z \in \mathcal{A}'' | [Z, A] = 0 \, \forall A \in \mathcal{A}''\}$, yet is almost disjoint from the microscopic observables: $\mathcal{C} \cap \mathcal{A} = \mathbb{C}\mathbf{1}$. Transfer of information to a macroscopic observable therefore implies perfect decoherence on all microscopic and macroscopic observables (cf. section III A 3).

Unfortunately, this transfer cannot be done by any automorphic time-evolution, since the macroscopic observables are central. Hepp proposed information transfer by a $t \to \infty$ limit of automorphisms. He was able to show that this causes decoherence in the weak-operator sense. That is, on each fixed microscopic observable, the coherence becomes arbitrarily small for sufficiently large t.

The paper was criticized by John Bell a few years later [Bel], on the grounds that, for each fixed time t, there are observables to be found on which coherence is not small. Since Bell was of the opinion that a 'wave packet reduction', even on closed systems, 'takes over from the Schrödinger equation', this was not to his satisfaction. He did agree however that these observables would become arbitrarily difficult to observe in practice for large t.

By considering large but finite closed systems subject to unitary time evolution, we hope to clarify the role that macroscopic systems play in making us mistake coherent superpositions for classical mixtures. It seems striking that the same, simple mathematics can also be used to understand why open systems do undergo decoherence as soon as they lose information.

[Arn] M. Arndt et al., *Wave–Particle Duality of C_{60}–Molecules*, Nature **401**, 680–682, (1999).
[Bel] J. Bell, *'On Wave Packet Reduction in the Coleman–Hepp Model'*, Helv. Phys. Acta **48**, 93–98, (1975).
[Fri] J. Friedman et al., *Quantum Superposition of Distinct Macroscopic States*, Nature **406**, 43–46, (2000).
[Hep] K. Hepp, *'Quantum Theory of Measurement and Macroscopic Observables'*, Helv. Phys. Acta **45**, 237–248, (1972).
[Jan] B. Janssens, *'Quantum Measurement, a Coherent Description'*, `arxiv.org/abs/quant-ph/0503009`, (2004).
[J&Z] E. Joos, H. Zeh, *'The Emergence of Classical Properties Through Interaction with the Environment'*, Z. Phys. B **59**, 223–243, (1985).
[Neu] J. von Neumann, *'Mathematische Grundlagen der Quantenmechanik'*, Springer-Verlag, (1932).
[Sch] E. Schrödinger, *'Die Gegenwartige Situation in der Quantenmechanik'*, Naturwissenschaften **23**, 807–812, 823–828, 844–849, (1935).
[Tak] M. Takesaki, *Theory of Operator Algebras I*, Springer-Verlag, (1979).
[Wal] H. van der Wal et al., *Quantum Superposition of Macroscopic Persistent–Current States*, Science **290**, 773–777, (2000).
[Wer] R. Werner, *'Quantum Information Theory – an Invitation'*, Springer Tracts in Modern Physics **173**, 14–57, (2001).
[Zur] W. Zurek, *'Environment-induced Superselection Rules'*, Phys. Rev. D **26**, 1862–1880, (1982).

Protected Subspaces in Quantum Information

Krzysztof Majgier[1], Hans Maassen[2], Karol Życzkowski[1,3]
[1] Instytut Fizyki im. Smoluchowskiego, Uniwersytet Jagielloński, ul. Reymonta 4, 30-059 Kraków, Poland
[2] Department of Mathematics, University of Nijmegen, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands and
[3] Centrum Fizyki Teoretycznej, Polska Akademia Nauk, Al. Lotników 32/44, 02-668 Warszawa, Poland
(Dated: September 1, 2009)

In certain situations the state of a quantum system, after transmission through a quantum channel, can be perfectly restored. This can be done by "coding" the state space of the system before transmission into a "protected" part of a larger state space, and by applying a proper "decoding" map afterwards. By a version of the Heisenberg Principle, which we prove, such a protected space must be "dark" in the sense that no information leaks out during the transmission. We explain the role of the Knill-Laflamme condition in relation to protection and darkness, and we analyze several degrees of protection, whether related to error correction, or to state restauration after a measurement. Recent results on higher rank numerical ranges of operators are used to construct examples. In particular, dark spaces are constructed for any map of rank 2, for a biased permutations channel and for certain separable maps acting on multipartite systems. Furthermore, error correction subspaces are provided for a class of tri-unitary noise models.

I. INTRODUCTION

We consider a quantum channel of finite dimension through which a quantum system in some state is sent. The output consists of another quantum state, and possibly some classical information. We are interested in the question to what extent the original quantum state can be recovered from that state and that information. In particular, we investigate if there are subspaces of the Hilbert space of the original system, on which the state can be perfectly restored.

In the literature a hierarchy of such spaces, which we shall call *protected subspaces* here, has been described. The strongest protection possible is provided in the case of a "decoherence free subspace" [1–4]. In this case the channel acts on the subspace as a isometric transformation. All we have to do in order to recover the state, is to rotate it back.

The next strongest form of protection occurs when the channel acts on the subspace as a random choice between isometries, whose image spaces are mutually orthogonal. Then by measuring along a suitable partition of the output Hilbert space, it can be inferred from the output state which isometry has occurred, so that it can be rotated back. This situation is characterized by the well-known Knill–Laflamme criterion, [5, 6] and the protected subspace in this case is usually called an *error correction subspace*.

The weakest form of protection is provided in yet a third situation, which was encountered in the context of quantum trajectories and the purification tendency of states along these paths [7]. In this case the deformation of the state is not caused by some given external device, but by the experimenter himself, who is performing a Kraus measurement [8]. Also in this case the "channel" acts as a random isometry, but the image spaces need not be orthogonal. It is now the *measurement outcome* (not the output state), that betrays to the experimenter which isometry has taken place. Using this information, he is able to undo the deformation of the component of the state that lies in the subspace considered.

It should be emphasized that the latter form of protection is far from a general error correction procedure. The experimenter only repairs the damage that he himself has incurred by his measurement.

Nevertheless, the above situations seem mathematically sufficiently similar to deserve study under a common title.

In all these three cases the experimenter learns nothing during the recovery operation about the component of the state inside our subspace. In this sense these subspaces can be considered "dark", and this darkness is essential for the protection of information. Our main result (Theorem 3) is concerned with the equivalence between protection and darkness, which is a consequence of Heisenberg's principle that no information on an unknown quantum state can be obtained without disturbing it (Corollary 2).

The question arises, for what channels protected subspaces are to be be found. We consider several examples in their Kraus decompositions. In each decomposition, we look for subspaces on which the channel acts as a multiple of an isometry, to be called a *homometry* here. Obviously, every (Kraus) operator A acts homometrically on a one-dimensional space $\mathbb{C}\psi$; its image $\mathbb{C}A\psi$ is another one-dimensional space, and the shrinking factor is $\sqrt{\langle A\psi, A\psi \rangle} = \|A\psi\|$. However, one-dimensional spaces are useless as coding spaces for quantum states. What we shall need, therefore, is the recent theory of higher rank numerical ranges [9, 10]. With the help of this we shall be able to construct several examples.

The paper is organized as follows. A brief review of basic concepts including channels and instruments is presented in section II. We discuss Heisenberg's principle in Section III. and prove our main Theorem, Theorem 3 in Section IV. In subsequent sections we analyze different forms of protected subspaces and compare their properties. In section V we review the notion of higher rank numerical range and quote some results on existence in the algebraic compression problem. Some examples of dark subspaces are presented in section VI, while an exemplary problem of finding an error correction code for a specific model of tri–unitary noise acting on a $3 \times K$ system is solved in section VII.

II. CHANNELS AND INSTRUMENTS

Let $\mathcal{H}$ be a finite-dimensional complex Hilbert space, and let $\mathcal{B}(\mathcal{H})$ denote the space of all linear operators on $\mathcal{H}$. We consider $\mathcal{H}$ as the space of pure states of some quantum system. By a *quantum operation* or *channel* on this system we mean a completely positive map $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ mapping the identity operator $\mathbb{1} = \mathbb{1}_{\mathcal{H}}$ to itself. The map Φ describes the operation "in the Heisenberg picture", i.e. as an action on observables. Its description "in the Schrödinger picture", i.e. as an action on density matrices ρ, is described by its adjoint Φ^*. The maps Φ and Φ^* are related by

$$\forall_\rho \forall_{X \in \mathcal{B}(\mathcal{H})} : \quad \mathrm{tr}\big(\Phi^*(\rho) X\big) = \mathrm{tr}\big(\rho \Phi(X)\big) \ .$$

We note that the property $\Phi(\mathbb{1}) = \mathbb{1}$, which we require for Φ, is equivalent to trace preservation by Φ^*:

$$\mathrm{tr}\big(\Phi^*(\rho)\big) = \mathrm{tr}\big(\Phi^*(\rho) \cdot \mathbb{1}\big) = \mathrm{tr}\big(\rho \cdot \Phi(\mathbb{1})\big) = \mathrm{tr}(\rho \cdot \mathbb{1}) = \mathrm{tr}(\rho) \ .$$

By Stinespring's theorem, every channel $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ can be written as

$$\Phi(X) = V^\dagger (X \otimes \mathbb{1}_{\mathcal{M}}) V \ , \tag{2.1}$$

where V is an isometry $\mathcal{H} \to \mathcal{H} \otimes \mathcal{M}$ for some auxiliary Hilbert space $\mathcal{M}$. The minimal dimension r of $\mathcal{M}$ admitting such a representation is called the *Choi rank* [11, 12] of Φ.

Any Stinespring representation of Φ naturally leads to a wider quantum operation

$$\Psi : \mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{M}) \to \mathcal{B}(\mathcal{H}) : X \otimes Y \mapsto V^\dagger (X \otimes Y) V \ , \tag{2.2}$$

which can be interpreted (in the Heisenberg picture) as the result of coupling the system to some *ancilla* having Hilbert space $\mathcal{M}$.

Thus Stinespring's representation (2.1) can be symbolically rendered as in Fig. 1.

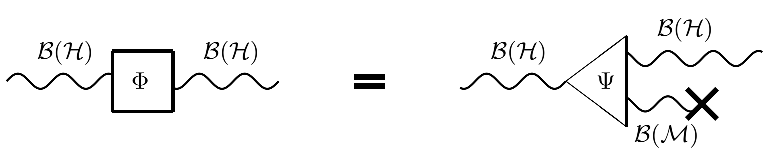


FIG. 1: Stinespring's dilation of Φ seen as coupling to an ancilla $\mathcal{M}$

In this picture, the cross stands for the substitution of $\mathbb{1}_{\mathcal{M}}$ (in the Heisenberg picture, reading from right to left), or the partial trace (in the Schrödinger picture, reading from left to right). Physically, it corresponds to throwing away, or just ignoring, the ancilla after the interaction. In the picture, the fact that Ψ is a *compression*, i.e. $\Psi = V^\dagger \cdot V$ for some isometry V, is symbolized by the triangular form of its box.

Now, by blocking the other exit in Fig. 1, we obtain the conjugate channel [13], Φ^C:

$$\Phi^C : \mathcal{B}(\mathcal{M}) \to \mathcal{B}(\mathcal{H}) : \quad Y \mapsto \Psi(\mathbb{1}_{\mathcal{H}} \otimes Y) = V^\dagger (\mathbb{1}_{\mathcal{H}} \otimes Y) V \ .$$

See also Fig. 2.

The main message of this paper is the following. The conjugate channel can be viewed as the flow of information into the environment. By Heisenberg's Principle, to be explained below, such a flow prohibits the faithful transmission of information through the original channel Φ. In particular, if the information encoded in some subspace of $\mathcal{H}$ is to be transmitted faithfully, nothing of it is visible from the outside: protection implies darkness. The degree of protection (decoherence free, strong or weak) is related to the degree of darkness, for which we shall define some terminology.

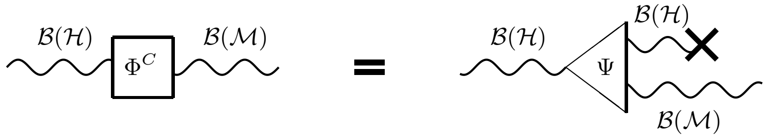


FIG. 2: The conjugate channel Φ^C.

Any orthonormal basis $f = (f_1, \ldots, f_m)$ in $\mathcal{M}$ corresponds to a possible von Neumann measurement Π_f^* on the ancilla, which maps a density matrix ρ on $\mathcal{M}$ to a probability distribution $(\langle f_1, \rho f_1 \rangle, \langle f_2, \rho f_2 \rangle, \ldots, \langle f_m, \rho f_m \rangle)$ on $\{1, 2, \ldots, m\}$. (Cf. Fig. 3.) In the Heisenberg picture this is the map from the algebra $\mathbb{C}^m$ with generators $e_1 = (1, 0, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$, ..., $e_m = (0, 0, \ldots, 0, 1)$, to $\mathcal{B}(\mathcal{M})$, given by

$$\Pi_f : \quad e_i \mapsto |f_i\rangle\langle f_i| \, .$$

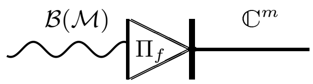


FIG. 3: Von Neumann measurement on $\mathcal{M}$.

In FIG. 3 the abelian algebra $\mathbb{C}^m$ is indicated by a straight line since it only carries classical information. Quantum information is designated by a wavy line.
Let us now denote by I_f the “partial inner product map”

$$\mathcal{H} \otimes \mathcal{M} \to \mathcal{H} : \quad \varphi \otimes \theta \mapsto \langle f, \theta \rangle \, \varphi \, ,$$

and let us write

$$A_i := I_{f_i} V \quad \in \mathcal{B}(\mathcal{H}) \, .$$

Then since $I_{f_i}^\dagger X I_{f_j} = X \otimes |f_i\rangle\langle f_j|$, we obtain a decomposition of Φ along the basis $(f_i)_{i=1}^m$ as follows:

$$\Phi(X) = \Psi(X \otimes \mathbb{1}_{\mathcal{M}}) = \sum_{i=1}^m \Psi\big(X \otimes |f_i\rangle\langle f_i|\big) = \sum_{i=1}^m V^\dagger I_{f_i}^\dagger X I_{f_i} V = \sum_{i=1}^m A_j^\dagger X A_j \, . \tag{2.3}$$

This is a *Kraus decomposition* of Φ. Combining the coupling to the ancilla with a von Neumann measurement on the latter, we obtain an *instrument* in the language of Davies and Lewis [14]:

$$\Psi_f : \mathcal{B}(\mathcal{H}) \otimes \mathbb{C}^m \to \mathcal{B}(\mathcal{M}) : \quad X \otimes e_i \mapsto V^\dagger (X \otimes |f_i\rangle\langle f_i|) V = A_i^\dagger X A_i \, . \tag{2.4}$$

The isometric property of V is now expressed as

$$V^\dagger V = \sum_{i=1}^m A_i^\dagger A_i = \mathbb{1} \, . \tag{2.5}$$

III. HEISENBERG'S PRINCIPLE OR OBSERVER EFFECT

In quantum mechanics observables are represented as self-adjoint operators on a Hilbert space. When A and B are commuting operators, then they possess a common complete orthonormal set of eigenvectors. Each of these eigenvectors ψ determines a state which associates sharply determined values to both observables A and B.

But when A and B do not commute, such states may not exist. This important property of quantum mechanics was first discussed by Heisenberg [15], and is called the *Heisenberg Uncertainty Principle*. It was formulated by Robertson [16] in the form

$$\sigma_\psi(A) \cdot \sigma_\psi(B) \geq \tfrac{1}{2} |\, \langle \psi, (AB - BA)\psi \rangle \,| \, .$$

Here $\sigma_\psi(X)$ is the standard deviation of X in the distribution induced by ψ. Already in the very same paper, Heisenberg introduced a second and very different principle, which is sometimes designated as the "Observer Effect", and which we shall call the Heisenberg Principle here. Roughly speaking, it says that:

if A and B do not commute,
a measurement of B perturbs the probability distribution of A. (3.1)

In the first half century of quantum mechanics, physicists, including Heisenberg himself, were satisfied with this formulation, and even considered it more or less identical to the Uncertainty Principle above.

In recent years it was realized that in fact we have here two different principles. Good quantitative formulations have been given of the Heisenberg Principle (for example [17, 18]). For the purpose of the present paper we are satisfied with a qualitative ('yes-or-no') version.

Let us first note that the formulation of the principle needs sharpening. As it stands, the condition is not needed: already in the trivial case that $A = B$ measurement of B changes the probability distribution of A. Indeed changing the probability distribution of an observable is the very purpose of measurement! And also, when A and B commute, but are correlated, then gaining information on B typically changes the distribution of A. A characteristic property of quantum theory only arises if we require that the outcome of the measurement of A is not used in the determination of the new probability distribution of B. Even then, some states may go through unchanged.

Corrected for these observations, the Heisenberg Principle reads:

For noncommuting A and B we cannot avoid that,
for some initial states, a measurement of B changes the distribution of A,
even if we ignore the outcome of the measurement. (3.2)

The contraposition of the statement turns out to be mathematically more tractible:

If the probability distribution of A is not *altered in any initial state*
— by us performing some measurement and ignoring its outcome —
then the object measured must commute with A. (3.3)

In this form it is sometimes called the 'nondemolition principle'.

Now let us make this statement precise. We start with a self-adjoint operator A on $\mathcal{H}$. Its distribution in the state ρ is determined by the numbers $\mathrm{tr}(\rho g(A))$ when g runs through the functions on the spectrum of A. Then some quantum operation is performed which on $\mathcal{B}(\mathcal{H})$ is described by a completely positive unit preserving map Φ. We require that for all states ρ and all functions f

$$\mathrm{tr}\big(\Phi^*(\rho)g(A)\big) = \mathrm{tr}\big(\rho g(A)\big) \ ,$$

which is equivalent to

$$\Phi\big(g(A)\big) = g(A) \ .$$

I.e.: all elements of the *-algebra $\mathcal{A}$ consisting of functions of A are left invariant by Φ. Let us denote the *commutant* of $\mathcal{A}$ by $\mathcal{A}'$,

$$\mathcal{A}' = \{X \in \mathcal{B}(\mathcal{H}) \,|\, \forall_{Y\in\mathcal{A}} \ : XY = YX\} \ . \tag{3.4}$$

Now, the quantum operation Φ is due to a measurement, so it is actually of the form

$$\Phi(X) = \Theta(X \otimes \mathbb{1}),$$

where $\Theta : \mathcal{B}(\mathcal{H}) \otimes \mathbb{C}^m \to \mathcal{B}(\mathcal{H})$ is some instrument whose outcomes, labeled $1, 2, \ldots, m$, in the state ρ have probabilities $p_1, p_2, \ldots, p_m$ to occur, where

$$p_j = \mathrm{tr}\big(\rho\Theta(\mathbb{1} \otimes e_j)\big) \ ,$$

and where $\mathrm{tr}\big(\rho\Theta(X \otimes e_j)\big)/p_j$ is the expectation of X, conditioned on the outcome j. (This situation is comparable to, but more general than, that of Ψ_f in (2.4).) Here $\mathbb{C}^m$ is the algebra of measurement outcomes. Generalizing to arbitrary $\mathcal{A}$, we may now formulate the Heisenberg Principle as follows.

Proposition 1 (Heisenberg Principle.) *Let $\mathcal{H}$ be a finite dimensional Hilbert space, and $\mathcal{B}$ some finite dimensional *-algebra. Let $\mathcal{A}$ be a sub-*-algebra of $\mathcal{B}(\mathcal{H})$, and let Θ be a completely positive unit preserving map $\mathcal{B}(\mathcal{H})\otimes\mathcal{B} \to \mathcal{B}(\mathcal{H})$. Suppose that for all $A \in \mathcal{A}$ we have*

$$\Theta(A \otimes \mathbb{1}) = A \; .$$

Then for all $B \in \mathcal{B}$

$$\Theta(\mathbb{1} \otimes B) \in \mathcal{A}' \; .$$

Proof: For any density matrix ρ on $\mathcal{H}$, define the quadratic form D_ρ on $\mathcal{B}(\mathcal{H}) \otimes \mathcal{B}$ by

$$D_\rho(X,Y) := \mathrm{tr}\rho\big(\Theta(X^*Y) - \Theta(X)^*\Theta(Y)\big) \; .$$

By the Cauchy-Schwartz inequality for the completely positive map Θ this quadratic form is positive semidefinite. By assumption we have for all $A \in \mathcal{A}$:

$$\begin{aligned} D_\rho(A\otimes\mathbb{1}, A\otimes\mathbb{1}) &= \mathrm{tr}\rho\big(\Theta(A^*A\otimes\mathbb{1}) - \Theta(A\otimes\mathbb{1})^*\Theta(A\otimes\mathbb{1})\big) \\ &= \mathrm{tr}\rho\big(A^*A\otimes\mathbb{1} - (A\otimes\mathbb{1})^*(A\otimes\mathbb{1})\big) = 0 \; . \end{aligned}$$

It then follows from the Cauchy-Schwartz inequality for D_ρ itself that $D_\rho(A\otimes\mathbb{1}, \mathbb{1}\otimes B) = 0$ for all $B \in \mathcal{B}$. But then

$$\begin{aligned} \mathrm{tr}\rho\big(A\Theta(\mathbb{1}\otimes B)\big) &= \mathrm{tr}\rho\big(\Theta(A\otimes\mathbb{1})\Theta(\mathbb{1}\otimes B)\big) = \mathrm{tr}\rho\big(\Theta\big((A\otimes\mathbb{1})(\mathbb{1}\otimes B)\big)\big) = \mathrm{tr}\rho\big(\Theta\big((\mathbb{1}\otimes B)(A\otimes\mathbb{1})\big)\big) \\ &= \mathrm{tr}\rho\big(\Theta(\mathbb{1}\otimes B)\Theta(A\otimes\mathbb{1})\big) = \mathrm{tr}\rho\big(\Theta(\mathbb{1}\otimes B)A\big). \end{aligned}$$

Since this holds for all ρ, it follows that $\Theta(\mathbb{1}\otimes B)$ commutes with A. □

By taking $\mathcal{A}$ and $\mathcal{B}$ abelian, say $\mathcal{A}$ generated by some observable A, and $\mathcal{B} = \mathbb{C}^m$ as above, and by choosing for Θ some instrument giving information about B, we obtain a statement of the type (3.3).

But there are other possible conclusions. We may choose $\mathcal{A} = \mathcal{B}(\mathcal{H})$, so that $\mathcal{A}' = \mathbb{C}\cdot\mathbb{1}_\mathcal{H}$. Then the Heisenberg principle says that, if we wish to make sure that any possible state ρ on $\mathcal{H}$ be unchanged by our measurement, no information at all concerning ρ can be gained. This is expressed by the following corollary and FIG. 4.

Corollary 2 *In the situation of Proposition 1, if for all $A \in \mathcal{B}(\mathcal{H})$ we have*

$$\Theta(A \otimes \mathbb{1}) = A \; ,$$

then there is a positive normalized linear form α on $\mathcal{B}$ such that for all $B \in \mathcal{B}$:

$$\Theta(\mathbb{1} \otimes B) = \alpha(B)\cdot\mathbb{1}_\mathcal{H} \; .$$

Indeed, the expectation of an outcome observable,

$$\mathrm{tr}(\Theta^*\rho)(\mathbb{1}\otimes B) = \mathrm{tr}(\rho\Theta(\mathbb{1}\otimes B)) = \mathrm{tr}(\rho\mathbb{1}_\mathcal{H})\cdot\mathrm{tr}(\alpha B) = \mathrm{tr}(\alpha B)$$

does not depend on ρ (see FIG. 4.)

IV. PROTECTION AND DARKNESS: THE KNILL-LAFLAMME CONDITION

Let $\mathcal{L}$ be a complex Hilbert space of dimension smaller than that of $\mathcal{H}$, and let $C : \mathcal{L} \to \mathcal{H}$ be some isometry. The range of C is a subspace of $\mathcal{H}$, isomorphic with $\mathcal{L}$. Let $\Gamma : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{L})$ denote the *compression map*

$$\Gamma(X) = C^\dagger X C \; .$$

Note that Γ is completely positive and identity-preserving. Compression maps are a convenient way of describing subspaces of a Hilbert space in the language of operations. Note that the operation Γ^* (in the Schödinger picture) embeds density matrices on $\mathcal{L}$ into the range of C:

$$\Gamma^*(\rho) = C\rho C^\dagger \; .$$

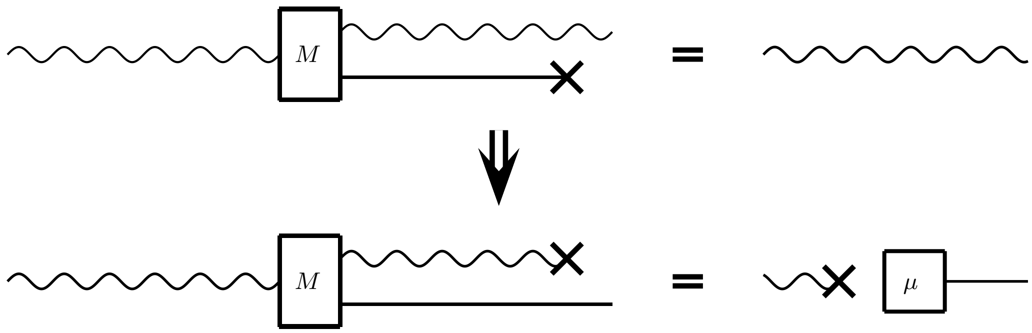


FIG. 4: Heisenberg's Principle as an implication between diagrams

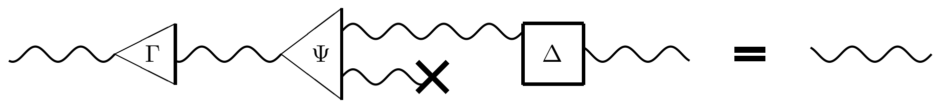


FIG. 5: Strong protection of Γ against Ψ

Physically, Γ is to be viewed as the "coding" operation.

Definition. We say that Γ (or the subspace $C\mathcal{L}$ of $\mathcal{H}$) is *protected against* a channel $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ if $\Gamma \circ \Phi$ is right-invertible, i.e. if there exists a "decoding" operation $\Delta : \mathcal{B}(\mathcal{L}) \to \mathcal{B}(\mathcal{H})$ such that

$$\Gamma \circ \Phi \circ \Delta = \mathrm{id}_{\mathcal{B}(\mathcal{L})} \ . \tag{4.1}$$

By virtue of (2.1) we may picture this state of affairs as in Fig. 5.
The subspace will be called *weakly protected against* an instrument $\Psi_f : \mathcal{B}(\mathcal{H}) \otimes \mathbb{C}^m \to \mathcal{B}(\mathcal{H})$ if $\Gamma \circ \Psi_f$ is right-invertible, i.e. if there exists a decoding operation $\Delta_f : \mathcal{B}(\mathcal{L}) \to \mathcal{B}(\mathcal{H}) \otimes \mathbb{C}^m$ such that

$$\Gamma \circ \Psi_f \circ \Delta_f = \mathrm{id}_{\mathcal{B}(\mathcal{L})} \ . \tag{4.2}$$

This is symbolically rendered in Fig. 6. The difference with Fig. 5 is that, in the case of weak protection, it is allowed to use the measurement outcome in the decoding. In the figure the classical information consisting of the measurement outcome, is symbolized by a straight line.

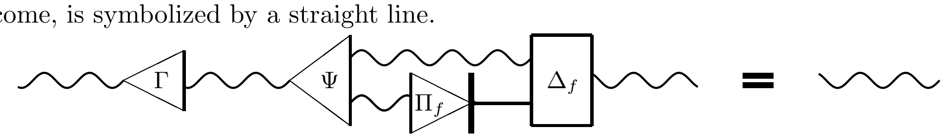


FIG. 6: Weak protection of Γ against Ψ_f

The above notions concern protection of information. Now we consider its availability to the external world.

Definition. Let $\Psi_f : \mathcal{B}(\mathcal{H}) \otimes \mathbb{C}^m \to \mathcal{B}(\mathcal{H})$ denote a quantum measurement (instrument) as described in (2.4). The subspace $C\mathcal{L} \subset \mathcal{H}$ (or the compression operation $\Gamma = C^\dagger \cdot C$), will be called *dark* with respect to Ψ_f if for all $i = 1, \ldots, m$ we have

$$\Gamma \circ \Psi_f(\mathbb{1} \otimes e_i) \in \mathbb{C} \cdot \mathbb{1}_{\mathcal{L}} \ . \tag{4.3}$$

This condition can be written in an equivalent form,

$$C^\dagger A_i^\dagger A_i C = \lambda_i \cdot \mathbb{1}_{\mathcal{L}} \quad \text{for} \quad i = 1, \ldots, m \ . \tag{4.4}$$

The subspace $C\mathcal{L}$ will be called *completely dark* for a channel $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ if it is dark for all Kraus measurements Ψ_f obtained by choosing different orthonormal bases in the ancilla space of some Stinespring dilation of Φ; i.e.

$$\forall_{Y \in \mathcal{B}(\mathcal{M})} : \quad \Gamma \circ \Psi(\mathbb{1} \otimes Y) \in \mathbb{C} \cdot \mathbb{1}_{\mathcal{L}} \ . \tag{4.5}$$

In terms of Kraus operators this is equivalent with the *Knill-Laflamme condition*:

$$C^\dagger A_i^\dagger A_j C = \alpha_{i,j} \cdot \mathbb{1}_{\mathcal{L}} \quad \text{for} \quad i, j = 1, \ldots, m \ . \tag{4.6}$$

Interpretation: From (4.3) and (4.4) we see that, if the von Neumann measurement along f is performed, the measurement outcome i has the same probability $\rho\big(\Gamma \circ \Psi_f(\mathbb{1} \otimes e_i)\big) = \rho(C^\dagger A_i^\dagger A_i C) = \lambda_i$, in all system states ρ, i.e. no information concerning the state ρ can be read off from the f-measurement on the ancilla.
Complete darkness (i.e. (4.5) or the equivalent Knill-Laflamme condition (4.6)) says that no information whatsoever concerning the input state reaches the ancilla. Mathematically, the Knill-Laflamme condition says that the range of the conjugate channel lies entirely in the center $\mathbb{C} \cdot \mathbb{1}_{\mathcal{L}}$ of $\mathcal{B}(\mathcal{L})$. Let us emphasize again that if the space C satisfies the conditions (4.6) for a map Ψ represented by a particular set of the Kraus operators $\{A_i\}_{i=1}^m$, then C also satisfies them for any other set of Kraus operators $\{B_i\}_{i=1}^{m'}$, used to represent the same map Ψ.
Note also that the set of conditions (4.6), which express complete darkness, naturally defines a state α, on the ancilla by a relation

$$\Gamma \circ \Psi(\mathbb{1} \otimes Y) = \mathrm{tr}(\alpha Y) \cdot \mathbb{1}_{\mathcal{L}} \ . \tag{4.7}$$

satisfied by any Y. This quantum state acting on an auxiliary system is called the *error correction matrix*, since the density matrix α_{ij} appears in eq. (4.6). Observe that the density operator α depends only on the map Ψ and not on the concrete form of the Kraus operators A_i, which represent the map and determine the matrix representation α_{ij} of α. Relations between matrix elements of the same state represented in two different basis are governed by the Schrödinger lemma [12], also called GHJW lemma [19, 20].

We are now going to prove the equivalence of protection and darkness. In the case of strong protection and complete darkness this reproduces and puts into perspective the result of Knill and Laflamme [6] In that case, if the state α is pure, then the decoding operation Δ can be realized by a unitary evolution, Hence the purity constraint for the error correction matrix, $\alpha = \alpha^2$, is the correct condition for a decoherence free subspace [21] – see also the proof of Theorem 3. As a quantitative measure, which characterizes to what extent a given protected space is close to a decoherence free space, one can use the von Neumann entropy of this state, $S = -\mathrm{Tr}\alpha \ln \alpha$. This *code entropy* [22] is equal to zero if the protected space is decoherence free or if the information lost can be recovered by a reversible unitary operation. Observe that the code entropy S characterizes the map Ψ and the code space C, but does not depend on the particular Kraus form used to represent Ψ.

In this way we have determined a hierarchy in the set of protected spaces. Every decoherence free subspace belongs to the class of completely dark subspaces, which correspond to error correction codes. In turn the completely dark subspaces form a subset of the set of dark subspaces – see Fig. 7.

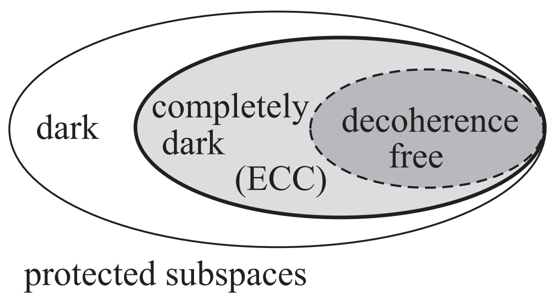


FIG. 7: Sketch of the hierarchy of protected subspaces.

Theorem 3 (Equivalence of Protection and Darkness) *Let $\mathcal{H}$, $\mathcal{M}$, and $\mathcal{L}$ be finite dimensional Hilbert spaces. Let $C : \mathcal{L} \to \mathcal{H}$ and $V : \mathcal{H} \to \mathcal{H} \otimes \mathcal{M}$ be isometries, and let Φ, Ψ and Ψ_f be as defined in (2.1), (2.2) and (2.4). Then $C\mathcal{L}$ is weakly protected against the instrument Ψ_f if and only if $C\mathcal{L}$ is dark for Ψ_f. It is strongly protected against Φ if and only if it is completely dark for Φ.*

Proof:

First assume that $C\mathcal{L}$ is strongly protected against Φ, i.e. (4.1) holds for some decoding operation Δ. Let $\Phi(X) = \Psi(X \otimes \mathbb{1})$ for some compression Ψ. Define

$$\Theta : \mathcal{B}(\mathcal{L}) \otimes \mathcal{B}(\mathcal{M}) \to \mathcal{B}(\mathcal{L}) : X \otimes Y \mapsto \Gamma \circ \Psi(\Delta(X) \otimes Y) \ .$$

Then $\Theta(X \otimes \mathbb{1}) = X$ for all $X \in \mathcal{B}(\mathcal{L})$, and by Corollary 2, since $\Delta(\mathbb{1}) = \mathbb{1}$,

$$\Gamma \circ \Psi(\mathbb{1} \otimes Y) = \Theta(\mathbb{1} \otimes Y) \in \mathbb{C} \cdot \mathbb{1} \ .$$

so (4.5) holds, and $C\mathcal{L}$ is completely dark for Φ.
Conversely, suppose that $C\mathcal{L}$ is completely dark for Ψ, and let α denote the density matrix given by (4.7) Then we may diagonalize:

$$\mathrm{tr}(\alpha Y) = \sum_{i=1}^{m} a_i \, \langle \, f_i, Y f_i \, \rangle$$

for some orthonormal set $(f_i)_{i=1}^{m'}$ (with $m' \leq m$) of $\mathcal{B}(\mathcal{M})$ and positive numbers $a_1, a_2, \ldots, a_{m'}$ summing up to 1. Now let $A_i := I_{f_i} V$. Then for all $\psi \in \mathcal{L}$:

$$\begin{aligned} \langle \, A_i C\psi, A_j C\psi \, \rangle &= \left\langle \, I_{f_i} V C\psi, I_{f_j} V C\psi \, \right\rangle \\ &= \left\langle \, \psi, C^\dagger V^\dagger (\mathbb{1} \otimes |f_i\rangle\langle f_j|) V C\psi \, \right\rangle \\ &= \alpha(|f_i\rangle\langle f_j|) \cdot \|\psi\|^2 \\ &= a_i \delta_{ij} \cdot \|\psi\|^2 \; . \end{aligned}$$

So the ranges of $A_i C$ and $A_j C$ are orthogonal for $i \neq j$ and A_i is homometric on $C\mathcal{L}$. Now define D_i for $i = 1, 2, \ldots, m'$ on these orthogonal ranges by

$$D_i \varphi = 0 \quad \text{if} \quad \varphi \perp \mathrm{Range}\,(A_i C), \qquad D_i A_i C \psi = \sqrt{a_i} \psi \; .$$

(D_i "rotates back" the action of $A_i C$.) Let Δ denote the operation

$$\Delta(Z) := \sum_{i=1}^{m'} D_i^\dagger Z D_i + \rho(Z) \left(\mathbb{1}_{\mathcal{H}} - \sum_{j=1}^{m'} D^\dagger D_j \right) \; .$$

for some arbitrary state ρ on $\mathcal{B}(\mathcal{L})$. (The term with ρ is intended to ensure that $\Delta(\mathbb{1}_{\mathcal{L}}) = \mathbb{1}_{\mathcal{H}}$.) Then we have for all $Z \in \mathcal{B}(\mathcal{L})$:

$$\begin{aligned} \Gamma \circ \Phi \circ \Delta(Z) &= \sum_{j=1}^{m'} \sum_{i=1}^{m'} C^\dagger A_j^\dagger D_i^\dagger Z D_i A_j C \\ &= \sum_{j=1}^{m'} \sum_{i=1}^{m'} \frac{1}{a_i} C^\dagger A_j^\dagger A_i C Z C^\dagger A_i^\dagger A_j C = \sum_{ij=1}^{m'} \delta_{ij} a_i Z = Z \; . \end{aligned}$$

So $C\mathcal{L}$ is strongly protected against Φ by (4.1).

Now let us prove the equivalence between weak protection and darkness. Assume that $C\mathcal{L}$ is weakly protected against Ψ_f, i.e. (4.2) holds for some $\Delta_f : \mathcal{B}(\mathcal{L}) \to \mathcal{B}(\mathcal{H}) \otimes \mathbb{C}^m$, say $\Delta_f(X) = \sum_{j=1}^{m} \Delta_f^j(X) \otimes e_j$. Define $\Theta : \mathcal{B}(\mathcal{H}) \otimes \mathbb{C}^m \to \mathcal{B}(\mathcal{H})$ by

$$\Theta(X \otimes g) := \sum_{j=1}^{m} g(j) \Gamma \circ \Psi_f(\Delta^j(X) \otimes e_j) \; .$$

Then by (4.2), $\Theta(X \otimes \mathbb{1}) = X$ for all $X \in \mathcal{B}(\mathcal{L})$. Hence by Corollary 2,

$$\Gamma \circ \Psi_f(\mathbb{1} \otimes e_i) = \Theta(\mathbb{1} \otimes e_i) \in \mathcal{B}(\mathcal{H})' = \mathbb{C} \cdot \mathbb{1}_{\mathcal{L}} \; .$$

So (4.3) holds, and $C\mathcal{L}$ is dark for Ψ_f.

Conversely, assuming that $C\mathcal{L}$ is dark for Ψ_f, then $A_l C$ is homometric on $\mathcal{L}$ by (4.4), and we may define $D_l : \mathcal{H} \to \mathcal{L}$ by

$$D_l A_l C \psi := \sqrt{\lambda_l} \psi \quad \text{if } \psi \in \mathcal{L}, \qquad D_l \varphi = 0 \quad \text{if } \varphi \perp \mathrm{Range}\,(A_l C) \; .$$

(Briefly: $D_l = C^\dagger A_l^\dagger / \sqrt{\lambda_l}$ if $\lambda_l \neq 0$, zero otherwise.) Define the decoding operation $\Delta_f : \mathcal{B}(\mathcal{L}) \to \mathcal{B}(\mathcal{H}) \otimes \mathbb{C}^m$ by

$$\Delta_f(Z) := \bigoplus_{l=1}^{m} \left(D_l^\dagger Z D_l + (\mathbb{1}_{\mathcal{H}} - D_l^\dagger D_l) \rho(Z) \right)$$

for some arbitrary state ρ on $\mathcal{B}(\mathcal{L})$. Then, for $Z \in \mathcal{B}(\mathcal{L})$:

$$\begin{aligned}
\Gamma \circ \Psi_f \circ \Delta_f(Z) &= \Gamma \circ \Psi_f \left(\sum_{l=1}^{m} \big(D_l^\dagger Z D_l + (\mathbb{1} - D_l^\dagger D_l)\rho(Z) \big) \otimes e_l \right) \\
&= C^\dagger V^\dagger \left(\sum_{l=1}^{m} \big(D_l^\dagger Z D_l + (\mathbb{1} - D_l^\dagger D_l)\rho(Z) \big) \otimes |f_l\rangle\langle f_l| \right) V C \\
&= \sum_{l=1}^{m} C^\dagger A_l^\dagger D_l^\dagger Z D_l A_l C = \sum_{l=1}^{m} \frac{1}{\lambda_l} (C^\dagger A_l^\dagger A_l C) Z (C^\dagger A_l^\dagger A_l C) = \sum_{l=1}^{m} \lambda_l Z = Z \ .
\end{aligned}$$

□

V. COMPRESSION PROBLEMS AND GENERALIZED NUMERICAL RANGE

For a given channel $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ we are interested in the protected subspaces of $\mathcal{H}$. These are the subspaces on which the compressions of $A_i^\dagger A_j$ act as scalars. In this section we review this compression problem.
Let T be an operator acting on a Hilbert space $\mathcal{H}$ of dimension n, say. For any $k \geq 1$, define the *rank-k numerical range* of T to be the subset of the complex plane given by

$$\Lambda_k(T) = \big\{ \lambda \in \mathbb{C} : C^\dagger T C = \lambda \mathbb{1} \text{ for some } C : \mathbb{C}^k \to \mathcal{H} \big\}, \tag{5.1}$$

The elements of $\Lambda_k(T)$ can be called "compression-values" for T, as they are obtained through compressions of T to a k-dimensional *compression subspace*. The case $k = 1$ yields the standard numerical range for operators [23]

$$\Lambda_1(T) = \{ \langle \psi | T \psi \rangle \ : \ |\psi\rangle \in \mathcal{H} \, , \langle \psi | \psi \rangle = 1 \}. \tag{5.2}$$

It is clear that

$$\Lambda_1(T) \supseteq \Lambda_2(T) \supseteq \ldots \supseteq \Lambda_n(T). \tag{5.3}$$

The sets $\Lambda_k(T)$, $k > 1$, are called *higher-rank numerical ranges* [9, 24]. For any normal operator acting on $\mathcal{H}_n$ this is a compact subset of the complex plane. For unitary operators this set is included inside every convex hull $(\mathrm{co}\,\Gamma)$, where Γ is an arbitrary $(n+1-k)$-point subset (counting multiplicities) of the spectrum of T [9]. It was recently shown that for any normal operator the sets $\Lambda_k(T)$ are convex [25, 26] while further properties of higher rank numerical range were investigated in [27–29].
The higher rank numerical range is easy to find for any Hermitian operator, $T = T^\dagger$ acting on an n-dimensional Hilbert space $\mathcal{H}$. Let us quote here a useful result proved in [9].

Lemma 4 *Let $x_1 \leq x_2 \leq \cdots \leq x_n$ denote the ordered spectrum (counting multiplicities) of a hermitian operator T. The rank-k numerical range of T is given by the interval*

$$\Lambda_k(T) = [x_k, x_{n+1-k}] \ , \tag{5.4}$$

Note that the higher rank numerical range of a hermitian T is nonempty for any $k \leq \mathrm{int}[(n+1)/2]$. Let us demonstrate an explicit construction of a compression to $\mathbb{C}^2$ which solves equation (5.1) for a Hermitian matrix T of size $n = 4$. The latter's eigenvalue equation reads $T|\phi_i\rangle = x_i|\phi_i\rangle$. Choose any real $\lambda \in \Lambda_2(T) = [x_2, x_3]$. It may be represented as a convex combination of two pairs of eigenvalues $\{x_1, x_3\}$ and $\{x_2, x_4\}$ – see Fig. 8a. Writing

$$\lambda \ = \ (1-a)x_1 + a x_3 \ = \ (1-b)x_2 + b x_4 \tag{5.5}$$

one obtains the weights

$$a = \frac{\lambda - x_1}{x_3 - x_1} =: \sin^2\theta_1 \quad \text{and} \quad b = \frac{\lambda - x_2}{x_4 - x_2} =: \sin^2\theta_2 \tag{5.6}$$

which determine real phases θ_1 and θ_2. These phases allow us to define an isometry $C : \mathbb{C}^2 \to \mathcal{H}$ by

$$C : \left\{ \begin{array}{lll} e_1 & \mapsto & \cos\theta_1 |\phi_1\rangle + \sin\theta_1 |\phi_3\rangle \\ e_2 & \mapsto & \cos\theta_2 |\phi_2\rangle + \sin\theta_2 |\phi_4\rangle \end{array} \right. \ , \tag{5.7}$$

Observe that

$$\left\langle\, e_1, C^\dagger T C e_1 \,\right\rangle = \cos\theta_1 x_1 \langle\phi_1|\psi_1\rangle + \sin\theta_1 x_3 \langle\phi_3|\psi_1\rangle = (1-a)x_1 + a x_3 = \lambda. \tag{5.8}$$

Similarly, we have $\left\langle\, e_2, C^\dagger T C e_2 \,\right\rangle = \lambda$. Further, we also have $\left\langle\, e_1, C^\dagger T C e_2 \,\right\rangle = 0 = \left\langle\, e_2, C^\dagger T C e_1 \,\right\rangle$. It follows that $C^\dagger T C = \mathbb{1}$, and the isometry (5.7) provides a solution of the compression problem (5.1) as claimed. Note that one can select another pairing of eigenvalues, and the choice $\{x_1, x_4\}$ and $\{x_2, x_3\}$ allows us to get in this way another subspace $C'\mathcal{L}$ spanned by vectors obtained by a superposition of states $|\phi_1\rangle$ with $|\phi_4\rangle$ and $|\phi_2\rangle$ with $|\phi_3\rangle$ respectively.

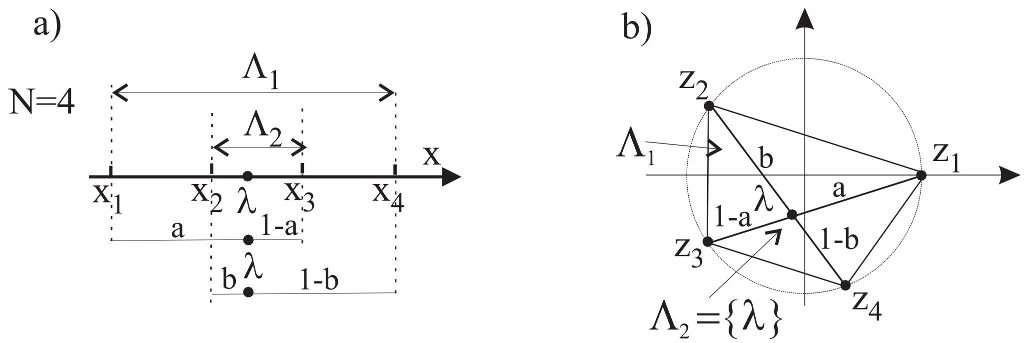


FIG. 8: Standard numerical range Λ_1 and higher rank numerical range Λ_2 for a) Hermitian operator T of size 4 and b) non–degenerate unitary $U \in U(4)$. Observe similarity in finding the weights a and b used to construct superposition of states forming the subspace $C\mathcal{L}$ in both problems.

For a given operator T one may try to solve its compression equation (5.1) and look for its numerical range $\Lambda_k(T)$. Alternatively, one may be interested in the following simple *compression problem*: For a given operator T find all possible subspaces $C\mathcal{L}$ of a fixed size k which satisfy (5.1).

Furthermore, it is natural to raise a more general, *joint compression problem* of order M. For a given set of M operators $\{T_1, \ldots T_M\}$ acting on $\mathcal{H}_n$ find a subspace $C\mathcal{L}$ of dimensionality k which solves simultaneously M compression problems:

$$C^\dagger T_m C = \lambda_m \mathbb{1} \quad \text{for} \quad m = 1, \ldots, M\ . \tag{5.9}$$

Note that all compression constants, $\lambda_m \in \Lambda_k(T_m)$, can be different, but the isometry C needs to be the same.

VI. DARK SUBSPACES

In this section we provide several results concerning existence of darks spaces for several classes of quantum maps.

A. Random external fields

Consider a noisy channel Φ given by

$$\Phi_U(X) = \sum_{i=1}^{r} q_i\, U_i^\dagger X U_i, \tag{6.1}$$

where all operators U_i are unitary while positive weights q_i sum up to unity. Such maps are called *random external fields* [30] or random unitary channels. The standard Kraus form (2.3) is obtained by setting $A_i = \sqrt{q_i} U_i$.

In this Kraus decomposition the whole space, and hence every subspace, is dark. This corresponds to the fact that the choice between the unitaries, which is made with the probability distribution $(q_1, \ldots, q_r)$, gives no information on the quantum state. And indeed, knowledge of the "external field", i.e. of the outcome i, permits us to undo, by the inverse of U_i, the action of the channel.

B. Rank two quantum channels

Let us now analyze a rank two channel,

$$\rho' = \Phi_2(\rho) = A_1 \rho A_1^\dagger + A_2 \rho A_2^\dagger \ , \tag{6.2}$$

Lemma 5 *For any Kraus representation of any rank-two channel acting on a system of size N there exist a dark subspace of dimension $k = \mathrm{int}[(N+1)/2]$.*

Proof. We need to solve a joint compression problem (5.9) of order two, for two Hermitian operators $T_1 = A_1^\dagger A_1$ and $T_2 = A_2^\dagger A_2$. Due to Lemma 4 there exists a subspace P_k of dimension $k = \mathrm{int}[(N+1)/2]$ which solves the compression problem for the Hermitian operator T_1 of size N. It is also a solution of the compression problem for the other operator, since the trace preserving condition implies $T_2 = \mathbb{1} - T_1$. □

C. Biased permutation channel

Consider a quantum map acting on a system of arbitrary size n described by the Kraus form (2.3). Let us assume that all Kraus operators are given by 'biased permutations'

$$A_i = P_i \sqrt{D_i} \ , \quad i = 1, \dots, r \ . \tag{6.3}$$

where D_i is a diagonal matrix containing non-negative entries, and P_i denotes an arbitrary permutation of the N-element set. Hence all elements of the POVM form diagonal matrices,

$$T_i = A_i^\dagger A_i = \sqrt{D_i} P_i^\dagger P_i \sqrt{D_i} = D_i \ , \tag{6.4}$$

in general not proportional to identity. Note that the Kraus operators defined in this way need not to be Hermitian. To satisfy the trace preserving condition (2.5) we need to assume that $\sum_{i=1}^r D_i = \mathbb{1}$. Let us define an auxiliary rectangular matrix of size $r \times N$, namely $S_{im} := (D_i)_{mm} \geq 0$. Then the above constraints for the matrices D_i is equivalent to the statement that S is *stochastic*, since the sum of all elements in each column is equal to 1,

$$\sum_{i=1}^r S_{im} = 1 \quad \text{for} \quad m = 1, \dots, N \ . \tag{6.5}$$

A map described by Kraus operators fulfilling relations (6.3) and (6.5) will be called a *biased permutation channel*.

We are going to construct a dark space for a wide class of such channels. For simplicity assume that the size of the system is even, $N = 2k$. Let us additionally assume that all elements in each row of B are ordered (increasingly or decreasingly) and that the matrix S enjoys a symmetry relation,

$$S_{i,m} + S_{i,n-m+1} = \text{const} =: \lambda_i \quad \text{for} \quad i = 1, \dots, r; \quad m = 1, \dots, k = n/2 \ . \tag{6.6}$$

Then the numbers λ_i can be defined by a sum of the entries in each row, $\lambda_i = \frac{2}{N} \sum_{m=1}^N S_{im}$.

Lemma 6 *Assume that a biased permutation channel acting on a system of size $N = 2k$ possesses the symmetry relation (6.6). Then it has a dark space of dimension $k = n/2$.*

Proof. We need to find a joint compression subspace for the set of r elements of POVM given by diagonal matrices D_i, with $i = 1, \dots, r$. Since these matrices commute, they have the same set of eigenvectors, denoted by $|v_m\rangle$, $m = 1, \dots, N$. Due to symmetry relation (6.6) we know that the barycenter of each spectrum, λ_i belongs to the higher rank numerical range, $\Lambda_k(D_i)$. Furthermore, this relation shows that (for any i) the number λ_i can be represented as a sum of two eigenvalues of D_i with the same weights, $\lambda_i = \frac{1}{2}(D_i)_{mm} + \frac{1}{2}(D_i)_{m'm'}$ with $m' = n+1-m$. By construction this property holds for all operators D_i, $i = 1, \dots r$. Hence the general construction of the higher order numerical range for Hermitian operators [10] implies that the subspace

$$C_k := \sum_{i=1}^k |\psi_i\rangle\langle\psi_i| \quad \text{where} \quad |\psi_i\rangle := \frac{1}{\sqrt{2}}(|v_i\rangle + |v_{1-i+N}\rangle) \tag{6.7}$$

fulfills the joint compression problem for all operators $T_i = D_i$, $i = 1, \dots r$. Hence this subspace is dark as advertised. □

To watch the above construction in action consider a three biased permutation channel acting on a two qubit system. Hence we set $r = 3$ and $N = 4$, and assume that five real weights satisfy $0 < a < b < x/2 < 1/2$ and $0 < c < d < x/2$. They can be used to define the channel by a stochastic matrix S

$$S = \begin{pmatrix} a & b & x-b & x-a \\ c & d & x-d & x-c \\ a' & b' & b'' & a'' \end{pmatrix} \ , \tag{6.8}$$

where $a' = 1 - a - c$, $b' = 1 - b - d$, $a'' = 1 - 2x + a + c$ and $b'' = 1 - 2x + b + d$. Note that this matrix satisfies the symmetry condition (6.6), the elements in each row are ordered, while mean weights in each row read $\lambda_1 = \lambda_2 = x/2$ and $\lambda_3 = 2(1 - x)$.

To complete the definition of the channel we need to specify three permutation matrice of size four. For instance let us choose $P_1 = P_{(1,2,3,4)}$, $P_2 = P_{(1,2),(3,4)}$ and $P_3 = P_{(1,4,3,2)}$, where according to the standard notion, the subscripts contain the permutation cycles. Then the biased permutation channel is defined by the three Kraus operators

$$A_1 = \begin{pmatrix} 0 & \sqrt{b} & 0 & 0 \\ 0 & 0 & \sqrt{x-b} & 0 \\ 0 & 0 & 0 & \sqrt{x-a} \\ \sqrt{a} & 0 & 0 & 0 \end{pmatrix} \ , \ A_2 = \begin{pmatrix} 0 & \sqrt{d} & 0 & 0 \\ \sqrt{c} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{x-c} \\ 0 & 0 & \sqrt{x-d} & 0 \end{pmatrix} \ , \ A_3 = \begin{pmatrix} 0 & 0 & 0 & \sqrt{a''} \\ \sqrt{a'} & 0 & 0 & 0 \\ 0 & \sqrt{b'} & 0 & 0 \\ 0 & 0 & \sqrt{b''} & 0 \end{pmatrix} \ , \tag{6.9}$$

which satisfy the trace preserving condition (2.5).

Since the barycenter λ_i of the spectrum of the POVM element $T_i = D_i$ (given by a row of matrix (6.8)), is placed symmetrically, in all three cases it can be represented by a convex combination of pairs of eigenvalues with weights equal to 1/2. Thus we define two pure states

$$|\psi_1\rangle := \frac{1}{\sqrt{2}}(|v_1\rangle + |v_4\rangle) \ , \quad |\psi_2\rangle := \frac{1}{\sqrt{2}}(|v_2\rangle + |v_3\rangle) \ , \tag{6.10}$$

and the two dimensional subspace spanned by them, $C = |\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|$. It is easy to verify that the subspace C satisfies $C^\dagger T_1 C = \lambda_1 \mathbb{1} = C^\dagger T_2 C$ while $C^\dagger T_3 C = \lambda_3 \mathbb{1}$ so this space is dark. Note that the subspace $C\mathcal{L}$ cannot be used to design an error correcting code since $C^\dagger A_1^\dagger A_2 C \notin \mathbb{C} \cdot \mathbb{1}$.

D. Composed systems and separable channels

Consider a bipartite system of size $n = n_A \times n_B$. A quantum operation Φ acting on this bipartite system is called *local*, if it has a tensor product structure, $\Phi = \Phi_A \otimes \Phi_B$, where both maps Φ_A and Φ_B are completely positive and preserve the identity. If for both individual operations, Φ_A and Φ_B, there exist protected subspaces C_k and Q_l respectively, then the product subspace $C_k \otimes Q_l$ of size kl is also a protected subspace for the composite map $\Phi_A \otimes \Phi_B$.

Similar protected subspaces of the product form can be constructed for a wider class of *separable maps* (see e.g. [12]),

$$\rho' = \Phi^*(\rho) = \sum_{i=1}^{r} (A_i \otimes B_i)\rho(A_i \otimes B_i)^\dagger. \tag{6.11}$$

Assume that a subspace $C_k \in \mathcal{H}_{N_A}$ is a solution of the joint compression problem for the set of r operators $A_i^\dagger A_i$, while a subspace $Q_l \in \mathcal{H}_{N_B}$ does the job for the set of r operators $B_i^\dagger B_i$. It is then easy to see that the product subspace $C_k \otimes Q_l$ of dimension kl is a dark subspace for the separable map (6.11).

It is straightforward to extend lemmas 3 and 4 for separable maps acting on composite systems and apply them to construct protected subspaces with a product structure. On the other hand, if for certain problems such product code subspace do not exist, one may still find a code subspace spanned by entangled states. Such a problem for the tri–unitary model is solved in following section.

VII. UNITARY NOISE AND ERROR CORRECTION CODES

In this section we are going to study multiunitary noise (6.1), also called random external fields, and look for existence of error correction codes, i.e. completely protected subspaces. In general the number r of unitary operators defining the channel can be arbitrary but we will restrict our attention to the cases in which this number is small.

A. Bi–unitary noise model

The case in which $r = 2$, referred to as *bi-unitary noise* was recently analyzed in [10, 24]. Let us rewrite the dynamics in the form

$$\rho' = \Phi^*(\rho) = qV_1^{\rho}V_1^{\dagger} + (1-q)V_2\rho V_2^{\dagger} \ . \tag{7.1}$$

and assume that we deal with the system of two qubits. Then both unitary matrices V_1 and V_2 belong to $U(4)$ while probability p belongs to $[0,1]$. The problem of finding the compression C for the above map is shown to be equivalent to the case

$$\rho'' = \Phi^*(\rho) = q\rho + (1-q)U\rho U^{\dagger} \tag{7.2}$$

where $U = V_1^{\dagger}V_2$.
Thus the error correction matrix α of size two defined by eq. (4.7) reads

$$\alpha \ = \ \begin{pmatrix} q & \sqrt{q(1-q)}\lambda \\ \sqrt{q(1-q)}\lambda^* & 1-q \end{pmatrix} \tag{7.3}$$

where λ is solution of the compression problem for U

$$C^{\dagger}UC = \lambda \cdot \mathbb{1} \ . \tag{7.4}$$

Thus to find the error correction space for the bi–unitary model it is sufficient to solve the compression equation for a single operator U. A solution exists for any unitary U [10], but for simplicity we will consider here the generic case if the spectrum of U is not degenerated. Assume that the phases these unimodular numbers $z_1, ..., z_4$ are ordered and that $|\psi_i\rangle$ denote the corresponding eigenvectors.
Let λ denote the intersection point between two chords of the unit circle, z_1z_3 and z_2z_4; compare Fig. 8b. This point can be represented as a convex combination of each pair of complex eigenvalues,

$$\lambda \ = \ (1-a)z_1 + az_3 \ = \ (1-b)z_2 + bz_4 \ , \tag{7.5}$$

where the non–negative weights read

$$a = \frac{\lambda - z_1}{z_3 - z_1} =: \sin^2\theta_1 \quad \text{and} \quad b = \frac{\lambda - z_2}{z_4 - z_2} =: \sin^2\theta_2 \tag{7.6}$$

and determine real phases θ_1 and θ_2. Note similarity with respect to the construction used in the Hermitian case, in which (5.5) represents a convex combination of points on the real axis. In an analogy with the reasoning performed for a hermitian T we define according to (5.7) an orthonormal pair of vectors $|\psi_1\rangle$ and $|\psi_2\rangle$ and define the associated isometry $C : e_j \mapsto \psi_j$. Since $\langle U\psi_1|\psi_1\rangle = \lambda = \langle U\psi_2|\psi_2\rangle$ and $\langle U\psi_1|\psi_2\rangle = 0 = \langle U\psi_2|\psi_1\rangle$ then $CUC = \lambda\mathbb{1}$. Therefore λ belongs to $\Lambda_2(U)$ as claimed and the range of C provides the error correction code for the bi-unitray noise (7.2) acting on a two-qubit system.

In the case of doubly degenerated spectrum of U the complex number λ is equal to the degenerated eigenvalue, so its radius, $|\lambda|$, is equal to unity. In this case the matrix α given in (4.6) represents a pure state, $\alpha = \alpha^2$, so the two–dimensional subspace spanned by both eigenvectors corresponding to the degenerated eigenvalues is *decoherence free*.

Bi–unitary noise model for higher dimensional systems was analyzed in [24]. It was shown in this work that for a generic U of size N there exists a code subspace of dimensionality $k = \text{int}[(N+2)/3]$. This result implies that for a system of m qubits and a generic U of size $N = 2^m$ there exists an error correction code supported on $m-2$ qubits. Furthermore, if $N = d^m$ and $d \geq 3$, there exists a code supported on $m-1$ quantum systems of size d.

B. Tri–unitary noise model

Consider now a model of noise described by three unitary operations acting on a bipartite, $N = 2 \times N_B$ system,

$$\rho' = \Phi^*(\rho) = q_1 V_1 \rho V_1^\dagger + q_2 V_2 \rho W_2^\dagger + (1 - q_1 - q_2) V_3 \rho V_3^\dagger \ . \tag{7.7}$$

Performing a unitary rotation in analogy to (7.2) we obtain an equivalent form

$$\rho'' = \Phi^*(\rho) = q_1 \rho + q_2 U_1 \rho U_1^\dagger + (1 - q_1 - q_2) U_2 \rho U_2^\dagger \ . \tag{7.8}$$

The model is thus characterized by two unitary matrices of size N, namely $U_1 = V_1^\dagger V_2$ and $U_2 = V_1^\dagger V_3$. and two weights q_1 and q_2, which we assume to be positive with their sum smaller than unity.

To find a simplest error correction code for this model one needs to find a two-dimensional subspace, which forms a joint solution of three compression problems

$$\begin{cases} C^\dagger U_1 C &= \lambda_{U_1} \mathbb{1} \\ C^\dagger U_2 C &= \lambda_{U_2} \mathbb{1} \\ C^\dagger W C &= \lambda_W \mathbb{1} \end{cases} , \tag{7.9}$$

where $W = U_1^\dagger U_2$. Each of the above three problems may be solved using the notion of the higher rank numerical range of a unitary matrix. However, for generic unitary matrices U_1 and U_2 of size 4 the corresponding compression subspaces do differ. Thus for a typical choice of the unitary matrices the tri–unitary noise model will not have an error correction code, for which it is required that the subspace C solves all three problems simultaneously.

There exist several examples of two commuting matrices U_1 and U_2 of size $N = 4$, such that they possess the same solution C of the compression problem. However, to assure that it coincides with the solution of the same problem for $W = U_1^\dagger U_2$, we will analyze an exemplary system of size $n = 2 \times 3$. Consider two unitary matrices of a tensor product form,

$$\begin{cases} U_1 &= U_A^\dagger \otimes U_B \\ U_2 &= U_A \otimes U_B \end{cases} \tag{7.10}$$

where

$$U_A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-i\alpha} & 0 \\ 0 & 0 & e^{i\alpha} \end{pmatrix} \quad \text{and} \quad U_B = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\xi} \end{pmatrix} \ . \tag{7.11}$$

Observe that U_1 and U_2 do commute, so they share the same set of eigenvectors. Assume that the phases satisfy $\alpha \in (\pi/2, \pi)$ and $\xi \in \big(0, \min\{\alpha, 2(\pi - \alpha)\}\big)$. Then the ordered spectra of both matrices read

$$U_1 = \mathrm{diag}\Big\{1, e^{i\xi}, e^{i\alpha}, e^{i(\alpha+\xi)}, e^{-i\alpha}, e^{i(\xi-\alpha)}\Big\}, \qquad U_2 = \mathrm{diag}\Big\{1, e^{i\xi}, e^{-i\alpha}, e^{i(\xi-\alpha)}, e^{i\alpha}, e^{i(\alpha+\xi)}\Big\}, \tag{7.12}$$

and differ only by the order of the eigenvalues. Both unitary matrices are represented in Fig. 9 in which z_i, $i = 1, \ldots, 6$ denote the ordered eigenvalues of U_1 while $|\varphi_i\rangle, i = 1, \ldots, 6$ are eigenvectors of this matrix. The same states form also the set of eigenvectors of U_2, but they correspond to other eigenvalues. Let z_i' denote the ordered eigenvalues of U_2. Then $|\varphi_3\rangle$ corresponds to $z_3' = z_5$ while $|\varphi_5\rangle$ corresponds to $z_5' = z_3$.

The third of the unitaries also has also a tensor product form,

$$W = U_1^\dagger U_2 = (U_A^\dagger \otimes U_B)^\dagger (U_A \otimes U_B) = U_A^2 \otimes \mathbb{1}_2. \tag{7.13}$$

Hence the spectrum of W, denoted by z_i'', consists of three pairs of doubly degenerated eigenvalues, $W = \mathrm{diag}\Big\{1, 1, e^{-2i\alpha}, e^{-2i\alpha}, e^{2i\alpha}, e^{2i\alpha}\Big\}$, see Fig. 10.

Numerical range of rank two for matrices U_1, U_2 and W is shown in the pictured as a gray region. Each point $\lambda \in \Lambda_2(U_1)$ offers a subspace C_2 which forms a solution of the first of three equations (7.9). However, the other two equations restrict further constraints for λ.

To construct an error correction code for the tri-unitary noise model we are going to follow the strategy used above for solving the compression problem: we split the Hilbert space into a direct sum of two subspaces of size three, and try to construct a single state in each subspace. More formally we define the subspace

$$C_2 = \sum_{i=1}^{2} |\psi_i\rangle\langle\psi_i| \tag{7.14}$$

where each state is obtained by a coherent superposition of three eigenstates of U_1,

$$\begin{cases} |\psi_1\rangle &= \sqrt{a_1}|\varphi_1\rangle + \sqrt{a_3}|\varphi_3\rangle + \sqrt{a_5}|\varphi_5\rangle \\ |\psi_2\rangle &= \sqrt{a_2}|\varphi_2\rangle + \sqrt{a_4}|\varphi_4\rangle + \sqrt{a_6}|\varphi_6\rangle \end{cases}. \tag{7.15}$$

Since the unitary operators U_i can be expressed as tensor product of diagonal matrices (e.g. $U_2 = U_A \otimes U_B$), their joint set of eigenvectors consits of product pure states only. On the other hand, the states $|\psi_1\rangle$ and $|\psi_2\rangle$ are by construction entangled.

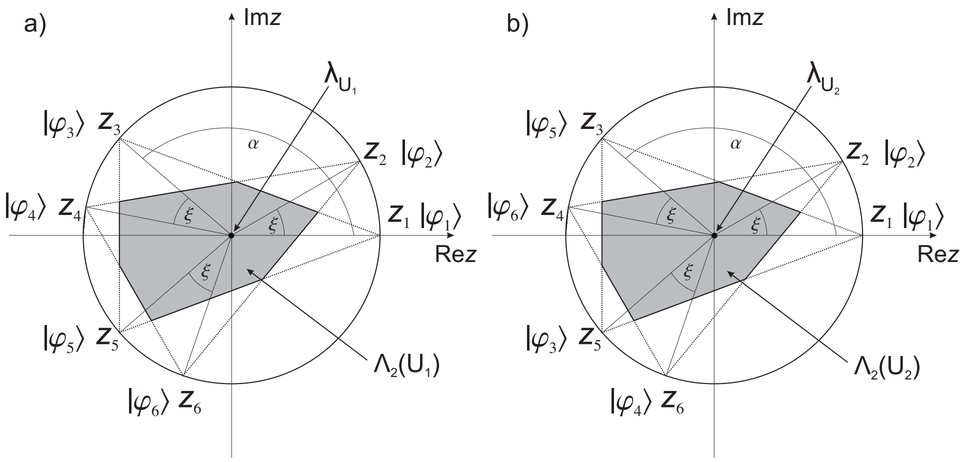


FIG. 9: Numerical range (gray space): a) $\Lambda_2(U_1)$; b) $\Lambda_2(U_2)$

The weights a_1 are defined as a weights obtained by representing point λ by a convex combination of the triples of eigenvalues. Since we wish to get a space C being a joint solution of all three equations (7.9), we are going to require that the same weights a_i can be used to form the compression value λ as a combination of both triples of eigenvalues for each spectrum,

$$\begin{cases} \lambda_{U_1} &= a_1 z_1 + a_3 z_3 + a_5 z_5 &= a_2 z_2 + a_4 z_4 + a_6 z_6 \\ \lambda_{U_2} &= a_1 z_1' + a_3 z_3' + a_5 z_5' &= a_2 z_2' + a_4 z_4' + a_6 z_6' \\ \lambda_W &= a_1 z_1'' + a_3 z_3'' + a_5 z_5'' &= a_2 z_2'' + a_4 z_4'' + a_6 z_6'' \end{cases} \tag{7.16}$$

where z_i, z_i' and z_i'' denote ordered spectra of U_1, U_2 and W, respectively. It is now clear that for a generic choice of U_1 and U_2 (which implies $W = U_1^\dagger U_2$), such a system has no solutions. However, if both diagonal matrices are of the special form (7.12), there exists a solution of the problem. The weights a_i satisfy

$$\begin{cases} a_1 &= a_2 &= 1 + \dfrac{1}{-1+\cos\alpha} \\ a_3 &= a_4 &= \dfrac{1}{2-2\cos\alpha} \\ a_5 &= a_6 &= \dfrac{1}{2-2\cos\alpha} \end{cases} \tag{7.17}$$

and imply the following compression values

$$\begin{cases} \lambda_{U_1} &= 0 \\ \lambda_{U_2} &= 0 \\ \lambda_W &= -1-2\cos\alpha \end{cases}. \tag{7.18}$$

Due to the symmetry of the problem the latter number λ_W is real.

Substituting the weights (7.17) into (7.15) we get an explicit form (7.14) of the subspace C. It is now easy to check that this subspace satisfies simultaneously all three equations (7.9) with compression values given by (7.18), hence it provides a two dimensional error correction code for this noise model. This solution is correct for any unitaries U_1 and U_2 having any set of eigenvectors $|\varphi_i\rangle$, $i = 1, \ldots, 6$ and spectra given by (7.12) and parameterized by phases α and ξ.

The above construction can be generalized for a tri–unitary noise model acting on larger system of size $N = 3 \times K$ [31]. An error correction code of size K exists in this case, if matrices U_1 and U_2 have the tensor product form (7.10), where $U_A = \mathrm{diag}\Big\{1, e^{i\alpha}, e^{-i\alpha}\Big\}$ as before and $U_B = \mathrm{diag}\Big\{1, e^{i\xi_2}, e^{i\xi_3}, \ldots, e^{i\xi_K}\Big\}$. The code subspace $C = \sum_{i=1}^{K} |\psi_i\rangle\langle\psi_i|$ is then obtained in an analogous way, by representing the Hilbert space as a direct product of K subspaces of dimension three each and constructing each state $|\psi_i\rangle$ as a coherent superposition of three eigenstates of U_1 corresponding to a triple of eigenvalues z_l, z_{l+K} and z_{l+2K} for $l = 1, \ldots K$. Note that the code space constructed here for the bipartite system does not have the tensor product structure, since it is spanned by entangled states (7.15).

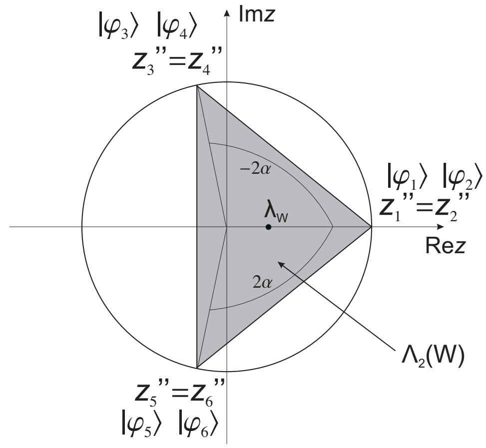


FIG. 10: Numerical range $\Lambda_2(W)$ is represented by a dark triangle

VIII. CONCLUSIONS

This paper concerns finite dimensional instruments or Kraus measurements, acting on a quantum system with Hilbert space $\mathcal{H}$. We have proved a version of Heisenberg's Principle, which connects 'darkness' to 'protection' of a subspace $\mathcal{L}$ of $\mathcal{H}$. 'Darkness' expresses the lack of visibility of the information contained in $\mathcal{L}$ from the measurement outcome, and 'protection' the degree to which this information remains present in the quantum system. Complete darkness corresponds to complete recoverability of information as in error correction codes.

We have presented examples of darkness and protection: instruments arising from random external fields, arbitrary rank 2 channels, and biased permutation channels. Bi-unitary noise models were analyzed recently in regard to their error correction properties in [10, 24]. Here we have also considered tri-unitary noise. For a a certain class of tri-unitary noise models acting on a $3 \times K$ quantum system, we have explicitly constructed an error correction code of size K. Although this particular noise model might be considered as not very realistic, we tend to believe that the technique proposed can be applied to a broader class of quantum systems.

IX. ACKNOWLEDGEMENTS

We enjoyed fruitful discussions with J. A. Holbrook, P. Horodecki and D. Kribs. We acknowledge financial support by the Polish Research Network LFPPI and by the European Research Project SCALA.

[1] L.-M. Duan and G.-C. Guo, *Phys. Rev. Lett.* **79**, 1953 (1997).
[2] P. Zanardi and M. Rasetti, *Phys. Rev. Lett.* **79**, 3306 (1997).
[3] D.A. Lidar, I.L. Chuang, and K.B. Whaley, *Phys. Rev. Lett.* **81**, 2594 (1998).
[4] A. Shabani and D. A. Lidar, *Phys. Rev.* **A 72**, 042303 (2005).
[5] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters *Phys. Rev.* **A 54**, 3824 (1996).
[6] E. Knill and R. Laflamme, *Phys. Rev.* **A 55**, 900 (1997).

[7] H. Maassen and B. Kümmerer, Purification of quantum trajectories. In: Institute of Mathematical Statistics, Lecture Notes – Monograph Series Vol. 48 (eds. Dee Denteneer, Frank den Hollander, Evgeny Verbitsky), pp. 252-261 (2006) and also quant-ph/0505084

[8] K. Kraus, General state changes in quantum theory, *Ann. Phys.* **64**, 311 (1971).

[9] M.-D. Choi, D. W. Kribs, and K. Życzkowski, Higher-Rank Numerical Ranges and Compression Problems, *Lin. Alg. Appl.* **418**, 828-839 (2006)

[10] M. D. Choi, D. W. Kribs and K. Życzkowski, Quantum error correcting codes from the compression formalism, *Rep. Math. Phys.* **58**, 77 (2006).

[11] M.-D. Choi, Completely positive linear maps on complex matrices, *Linear Alg. Appl.* **10**, 285 (1975).

[12] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement*, Cambridge University Press, Cambridge 2006.

[13] C. King, K. Matsumoto, M. Nathanson, M. B. Ruskai, Properties of Conjugate Channels with Applications to Additivity and Multiplicativity, *Markov Process Related Fields* **13**, 391-423 (2007).

[14] E. B. Davies, J. T. Lewis, An Operational Approach to Quantum Probability, *Comm. Math. Phys.* **17**, 239-260 (1970).

[15] W. Heisenberg, Über den anschaulichen Inhalt der Quantentheoretischen Kinematik und Mechanik, *Z. Phys.* **43**, 172-198 (1927).

[16] H. Robertson, The uncertainty principle, *Phys. Rev.* **34**, 163-164 (1929).

[17] R.F. Werner, The uncertainty relation for joint measurement of position and momentum, *Quant. Inf. Comp.* **4**, 546-562 (2004).

[18] B. Janssens, Unifying decoherence and the Heisenberg principle, `www.arxiv.org/quant-ph/0606093`.

[19] N. Gisin, Stochastic quantum dynamics and relativity, *Helv. Phys. Acta* **62**, 363 (1989).

[20] L. P. Hughston and R. Jozsa and W. K. Wootters, A complete classification of quantum ensembles having a given density matrix, *Phys. Lett.* **A 183**, 14 (1993).

[21] D.A. Lidar, D. Bacon, and K.B. Whaley, *Phys. Rev. Lett.* **82**, 4556 (1999).

[22] D.W. Kribs, A. Pasieka, K. Życzkowski, Entropy of a quantum error correction code, *Open Syst. Inf. Dyn.* **15**, 329-343 (2008)

[23] R. Bhatia, *Matrix Analysis*, Springer Verlag, New York 1997.

[24] M. D. Choi, J. A. Holbrook, D. W. Kribs and K. Życzkowski, *Operators and Matrices* **1**, 409 (2007).

[25] C.-K. Li, and N.-S. Sze, Canonical forms, higher rank numerical ranges, totally isotropic subspaces, and matrix equations, *Proc. Amer. Math. Soc.* **136**, 3013-3023 (2008).

[26] H. Woerdeman, The higher rank numerical range is convex, *Lin. & Multilin. Algebra* **56**, 65-67 (2008).

[27] M.-D. Choi, M. Giesinger J.A. Holbrook, D.W. Kribs, Geometry of higher-rank numerical ranges, *Lin. & Multilin. Algebra* **56**, 53-64 (2008).

[28] C.-K. Li, Y.-T. Poon and N.-S. Sze, Condition for the higher rank numerical range to be non-empty, *Lin. & Multilin. Algebra* **57**, 365-368 (2009).

[29] C.-K. Li, Y.-T. Poon and N.-S. Sze, Higher rank numerical ranges and low rank perturbations of quantum channels, *J. Math. Analysis Appl.* **348**, 843-855 (2008)

[30] R. Alicki and K. Lendi, *Quantum Dynamical Semigroups and Their Applications*, LNP 286, Springer, Berlin (1987)

[31] K. Majgier, Quantum error correction codes for unitary models of noise (*in Polish*), Master thesis, Jagiellonian University, Cracow, June 2007; see http://chaos.if.uj.edu.pl/∼karol/prace/Majgier07.pdf